

University of Windsor

Scholarship at UWindor

Electronic Theses and Dissertations

Theses, Dissertations, and Major Papers

2010

Artificial neural network and its applications in quality process control, document recognition and biomedical imaging

Mohammed Jahirul Islam
University of Windsor

Follow this and additional works at: <https://scholar.uwindsor.ca/etd>

Recommended Citation

Islam, Mohammed Jahirul, "Artificial neural network and its applications in quality process control, document recognition and biomedical imaging" (2010). *Electronic Theses and Dissertations*. 7972. <https://scholar.uwindsor.ca/etd/7972>

This online database contains the full-text of PhD dissertations and Masters' theses of University of Windsor students from 1954 forward. These documents are made available for personal study and research purposes only, in accordance with the Canadian Copyright Act and the Creative Commons license—CC BY-NC-ND (Attribution, Non-Commercial, No Derivative Works). Under this license, works must always be attributed to the copyright holder (original author), cannot be used for any commercial purposes, and may not be altered. Any other use would require the permission of the copyright holder. Students may inquire about withdrawing their dissertation and/or thesis from this database. For additional inquiries, please contact the repository administrator via email (scholarship@uwindsor.ca) or by telephone at 519-253-3000ext. 3208.

Artificial Neural Network and Its Applications in Quality Process Control, Document Recognition and Biomedical Imaging

by

Mohammed Jahirul Islam

A Dissertation
Submitted to the Faculty of Graduate Studies
through Electrical and Computer Engineering
in Partial Fulfillment of the Requirements for
the Degree of Doctor of Philosophy at the
University of Windsor

Windsor, Ontario, Canada
2010



Library and Archives
Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file *Votre référence*
ISBN: 978-0-494-62766-2
Our file *Notre référence*
ISBN: 978-0-494-62766-2

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.


Canada

© 2010 Mohammed Jahirul Islam

All Rights Reserved. No Part of this document may be reproduced, stored or otherwise retained in a retrieval system or transmitted in any form, on any medium by any means without prior written permission of the author.

Author's Declaration of Originality

I hereby certify that I am the sole author of this thesis and that no part of this thesis has been published or submitted for publication.

I certify that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my thesis and have included copies of such copyright clearances to my appendix.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution.

Abstract

In computer-vision based system a digital image obtained by a digital camera would usually have 24-bit color image. The analysis of an image with that many levels might require complicated image processing techniques and higher computational costs. But in real-time application, where a part has to be inspected within a few milliseconds, either we have to reduce the image to a more manageable number of gray levels, usually two levels (binary image), and at the same time retain all necessary features of the original image or develop a complicated technique. A binary image can be obtained by thresholding the original image into two levels. Therefore, thresholding of a given image into binary image is a necessary step for most image analysis and recognition techniques. In this thesis, we have studied the effectiveness of using artificial neural network (ANN) in pharmaceutical, document recognition and biomedical imaging applications for image thresholding and classification purposes. Finally, we have developed edge-based, ANN-based and region-growing based image thresholding techniques to extract low contrast objects of interest and classify them into respective classes in those applications.

Real-time quality inspection of gelatin capsules in pharmaceutical applications is an important issue from the point of view of industry's productivity and competitiveness. Computer vision-based automatic quality inspection and controller system is one of the solutions to this problem. Machine vision systems provide quality control and real-time feedback for industrial processes, overcoming physical limitations and subjective judgment of humans. In this thesis, we have developed an image processing system using edge-based image thresholding techniques for quality inspection that satisfy the industrial requirements in pharmaceutical applications to pass the accepted and rejected capsules.

In document recognition application, success of OCR mostly depends on the quality of the thresh-

olded image. Non-uniform illumination, low contrast and complex background make it challenging in this application. In this thesis, optimal parameters for ANN-based local thresholding approach for gray scale composite document image with non-uniform background is proposed. An exhaustive search was conducted to select the optimal features and found that pixel value, mean and entropy are the most significant features at window size 3x3 in this application. For other applications, it might be different, but the procedure to find the optimal parameters is same. The average recognition rate 99.25% shows that the proposed 3 features at window size 3x3 are optimal in terms of recognition rate and PSNR compare to the ANN-based thresholding technique with different parameters presented in the literature.

In biomedical imaging application, breast cancer continues to be a public health problem. In this thesis we presented a computer aided diagnosis (CAD) system for mass detection and classification in digitized mammograms, which performs mass detection on regions of interest (ROI) followed by the benign-malignant classification on detected masses. Three layers ANN with seven features is proposed for classifying the marked regions into benign and malignant and 90.91% sensitivity and 83.87% specificity is achieved that is very much promising compare to the radiologist's sensitivity 75%.

Acknowledgments

First of all, I would like to express my sincere appreciation to my co-supervisors, Dr. Majid Ahmadi and Dr. Maher A. Sid-Ahmed for their overall enthusiasm and devotion to their work that impressed me a lot. They are highly critical, while remaining very supportive. They not only provide invaluable guidance to my study, but also sets up the best role model for me and my colleagues to follow.

I also would like to thank my committee members, Dr. Roman Maev, Dr. Jonathan Wu and Dr. X. Chen, who monitor my work and take effort in reading and providing me with valuable comments on my thesis.

I was lucky to work with so many energetic and intelligent classmates and colleagues. I enjoyed every minutes I spend with them and cherish all of the help from them.

Especially, I would like to give my thanks to my parent, Mohammed Abdul Jalil, Rokeya Begum, my wife, Tahmina Haque Khan and my angels Zafir Raeid and Rahil Raeid. Without their continuous encourage and support, it is impossible for me to finish all of the jobs in the thesis.

Contents

Author's Declaration of Originality	iv
Abstract	v
Acknowledgments	vii
List of Figures	xii
List of Tables	xvi
List of Abbreviations	xviii
1 Introduction	1
1.1 Introduction	1
1.2 Image Segmentation	2
1.3 Quality Process Control of Pharmaceutical Manufactured Products	3
1.4 Document Image Analysis	3
1.5 Biomedical Imaging- Breast Cancer Detection	4
1.6 Thesis Objectives	5
1.7 Thesis Organization	6
2 Introduction to Artificial Neural Network	7
2.1 Introduction	7
2.2 Types of ANNs	8
2.2.1 Feed-forward Neural Network	8
2.2.2 Feedback Neural Network	8
2.3 The Learning Process of ANNs	10

2.3.1	Supervised Learning	10
2.3.2	Unsupervised Learning	10
2.4	Multi-layer Perceptron (MLP)	10
2.5	Advantages and Disadvantages of ANNs	12
2.5.1	Advantages of ANNs [2]	12
2.5.2	Disadvantages of ANNs [2]	12
2.6	Application of ANNs	13
2.7	Conclusions	13
3	Review of Image Thresholding Methods and Statistical Texture Features	15
3.1	Introduction	15
3.2	Survey Over Thresholding Techniques	16
3.2.1	Histogram-based Methods	16
3.2.2	Edge-based Methods	18
3.2.3	Clustering-based Methods	21
3.2.4	Entropy-based Methods	23
3.2.5	Local Adaptive Thresholding	23
3.2.6	Artificial Neural Network-based Methods	24
3.3	Performance Comparison of the Thresholding Techniques	25
3.4	Statistical and Texture Features	26
3.4.1	Pixel Value ¹	26
3.4.2	Mean ²	26
3.4.3	Standard Deviation ³	27
3.4.4	Smoothness ⁴	27
3.4.5	Entropy ⁵	27
3.4.6	Skewness ⁶	27
3.4.7	Kurtosis ⁷	28
3.4.8	Uniformity ⁸	28
3.5	ANN-based Thresholding Using Statistical Texture Features	28
3.6	Conclusion	29
4	Efficient ANN-based Thresholding in Document Recognition Application	32
4.1	Introduction	32
4.2	ANN-based Thresholding- Observations and Criticism	32

4.3	Proposed Method	37
4.3.1	Phase I-Optimal Parameters Selection	37
4.3.2	Phase II-Efficient ANN-based Thresholding Method	39
4.4	Simulation Results	40
4.4.1	Image Database	40
4.4.2	Training and Testing Results	41
4.4.3	Optimal Features Validation	43
4.4.4	Performance Evaluation of the Proposed Method	56
4.5	Conclusions	57
5	Edge-based Thresholding and Quality Process Control	62
5.1	Introduction	62
5.2	Existing Quality Control and Criticisms	63
5.3	Current State-of-the-art Solutions	65
5.4	Capsule Inspection System- OptiSorter	65
5.5	Proposed Methodology	66
5.5.1	Preprocessing	67
5.5.2	Image Segmentation	69
5.5.3	Border Tracing	71
5.5.4	Circle Fitting	72
5.6	Simulation Results and Performance Evaluation	73
5.6.1	Rationale Using Four Cameras	82
5.7	Conclusion	82
6	ANN-based Mass Classification in Biomedical Imaging	85
6.1	Introduction	85
6.2	Literature Survey	86
6.3	Haralick Texture Features [98, 105]	88
6.3.1	Gray Level Co-occurrence Matrix	88
6.3.2	Statistical Properties of GLCM and Haralick Features Calculation	89
6.4	Proposed Methods	91
6.4.1	ROI Preprocessing	92
6.4.2	Mass Extraction	95
6.4.3	Mass Classification	99

6.5	Simulation Results and Performance Evaluation	102
6.5.1	Image Database	102
6.5.2	Results and Performance	102
6.6	Validation	108
6.7	Conclusion	108
7	Conclusion and Future Research	114
7.1	Conclusions	114
7.2	Contributions and Future work	116
	References	118
A	Recognition Rate- Grid-based Method	125
	VITA AUCTORIS	129

List of Figures

1.1	Principal stages of image analysis system	2
2.1	Supervised learning of ANN	8
2.2	Multi-layer feed-forward network model	9
2.3	Complicated feed-forward/ feedback network	9
2.4	Flow chart of the training steps of MLP ANN	11
3.1	Sample Training Data Preparation	29
3.2	Sample Binary Image using NN-based Thresholding	30
4.1	Principal Stages of Document Analysis System	33
4.2	Sample Document Image- (poor contrast)	33
4.3	Sample Document Image- (non-uniform illumination)	34
4.4	Sample Document Image- (composite with complex background)	34
4.5	Sample Document Image- (Insect)	35
4.6	More Sample Document Images 1	35
4.7	More Sample Document Images 2	36
4.8	No. of features subset for each combination	38
4.9	Grid-based Method for Tabulating Recognition Rate- Subset vs. Window Size	39
4.10	Test Setup	39
4.11	Window Size Vs. Process Time (μ s) for each pixel	45
4.12	Window Size Vs. Character recognition rate (%)	46
4.13	Binary Image Using Proposed Method of figure 4.2	46
4.14	Binary Image Using Proposed Method of figure 4.3	47
4.15	Binary Image Using Proposed Method of figure 4.4	47

4.16 Sample Base Document 1	48
4.17 Sample Base Document 2	49
4.18 Sample Watermark 1	50
4.19 Sample Watermark 2	50
4.20 Sample Watermark 3	51
4.21 Sample Watermark 4	51
4.22 Sample Document Watermarked Image 1	52
4.23 Sample Document Watermarked Image 2	52
4.24 Sample Document Watermarked Image 3	53
4.25 Sample Document Watermarked Image 4	53
4.26 Binary Image 1 Using Proposed 3 Features	54
4.27 Binary Image 2 Using Proposed 3 Features	54
4.28 Binary Image 3 Using Proposed 3 Features	55
4.29 Binary Image 4 Using Proposed 3 Features	55
4.30 OCR output: From Binary Image 4.15 (Proposed) vs. Expected	57
4.31 OCR output: From Original Image vs. Expected	58
4.32 OCR output: From Binary Image 4.13 (Proposed) vs. Expected	58
4.33 OCR output: From Binary Image 4.14 (Proposed) vs. Expected	59
4.34 OCR output: From Binary Image (Using Otsu [37])	59
4.35 OCR output: From Binary Image (Using Niblack [54])	60
4.36 Segmentation Performance Comparison: Proposed, HMM[32], Kittler [38] and Otsu [37]	61
5.1 Typical defects (Source: Pharmaphil Inc., Windsor, ON)	63
5.2 Current capsule sorting system	64
5.3 Current capsule sorting system- Conveyor belt	64
5.4 Image acquisition system	66
5.5 Capsule image- Top, Bottom, Right and Left camera view	67
5.6 Typical capsule image (Source: Pharmaphill Inc.)	68
5.7 Block diagram of the proposed image processing system	68
5.8 Flow chart of the proposed capsule extraction method	71
5.9 Image enhancement using different Techniques	74
5.10 Quantitative measures of different image enhancement techniques	75
5.11 Image enhancement using Median filtering	76

5.12 Binary image using Otsu [37] and Kittler [38]	76
5.13 Training dataset generation using [67]	77
5.14 Binary image using ANN [67]	77
5.15 Edge-based segmentation	78
5.16 Closed view of edge-based segmentation	79
5.17 Good capsule image using the proposed method	79
5.18 Dented capsule image using the proposed method- Example 1	80
5.19 Dented capsule image using the proposed method- Example 2	80
5.20 Binary image of left camera image using proposed method	81
5.21 Binary image of right camera image using proposed method	82
5.22 Bottom and left camera view of a defective capsule	83
5.23 Right and top camera view of a defective capsule	83
6.1 Sample mammogram and principal stages of breast cancer detection	86
6.2 Example of GLCM for $d = 1$: (a) The original image (b) Final result of the horizontal GLCM $d = 1$. c) Four directions (0° , 45° , 90° , 135°)	89
6.3 X-ray label removal from the original mammogram	92
6.4 Breast extraction- Global ROI	93
6.5 ROI extraction and sample ROI	95
6.6 Sample Original and Enhanced ROIs	96
6.7 ROI division into blocks and seed block detection	97
6.8 Haralick Texture Features (1-4)	98
6.9 Haralick Texture Features (5-8)	98
6.10 Haralick Texture Features (9-13)	99
6.11 Flow chart of the proposed automatic seeded region growing method	100
6.12 Extracted mass from the ROI	100
6.13 Training data preparation and mass classification	101
6.14 Sample extracted mass using proposed method	103
6.15 Sample extracted contour using proposed method	103
6.16 Segmentation Performance Evaluation- Example 1	105
6.17 Segmentation Performance Evaluation- Example 2	105
6.18 Mass Classification- Example 1, Correct Classification (Benign mass)	106
6.19 Mass Classification- Example 2, Correct Classification (Malignant mass)	106

6.20 Mass Classification- Example 3, Misclassification (Classified as Benign, actually Malignant)	107
6.21 Validation- Example 1, Background, Object, Combined Image and Enhanced Image	109
6.22 Validation- Example 1, Automated Seed Selection	109
6.23 Validation- Example 1, Segmented Leaf, Extracted Contour and Extracted Leaf . . .	110
6.24 Validation- Example 2, Background, Object, Combined Image and Enhanced Image	110
6.25 Validation- Example 2, Automated Seed Selection	111
6.26 Validation- Example 2, Segmented Leaf, Extracted Contour and Extracted Leaf . . .	111
6.27 Validation- Example 3, Background, Object, Combined Image and Enhanced Image	112
6.28 Validation- Example 3, Automated Seed Selection	112
6.29 Validation- Example 3, Segmented Leaf, Extracted Contour and Extracted Leaf . . .	113
A.1 Recognition rate in various window size- Table A.1	126
A.2 Recognition rate in various window size- Table A.2	127
A.3 Recognition rate in various window size- Table A.3	128

List of Tables

3.1	Horizontal mask of Sobel Edge Detector	19
3.2	Vertical mask of Sobel Edge Detector	19
3.3	Horizontal mask of Prewit Edge Detector	19
3.4	Vertical mask of Prewit Edge Detector	19
3.5	Roberts' Edge Detector (-45°)	19
3.6	Roberts' Edge Detector ($+45^\circ$)	19
3.7	Neural Network Structure used by Alginahi [7]	30
4.1	Proposed Efficient ANN Structure	40
4.2	Comparative Statement of Recognition Rates- Using Binary and Original Image . .	42
4.3	Comparative Statement of Recognition Rates-Features Combination	42
4.4	Window Size Vs. Thresholding Time (Sec)	43
4.5	Window Size Vs. Correct Recognition Rate (%)	44
4.6	Window Size Vs. Process Time for each pixel	45
4.7	Number of Features Vs. Process Time (Sec)	45
4.8	PSNR Measurement for Different Parameters	48
4.9	Recognition Rate Comparison- Proposed, Niblack [54] and Otsu [37]- (Figure 4.4) . .	56
4.10	Performance Evaluation of the Proposed Method in Different Categories	57
4.11	Performance of the Proposed Method Using Optimal Features	59
5.1	Co-ordinates of eight neighbors of a pixel	72
5.2	Neural Network Architecture	78
5.3	Quality Inspection Result	84
5.4	Performance Evaluation	84

6.1	Simulation Result: Mass Segmentation	102
6.2	Proposed ANN Structure for Mass Classification	106
6.3	Comparative Statement of Classification	107
6.4	Comparative Statement of Sensitivity of the Proposed System to the Radiologist . .	107

List of Abbreviations

2-D	Two Dimensional.
AD	Area Difference.
ANN	Artificial Neural Network.
CAD	Computer Aided Diagnosis.
CCD	Charged Coupled Device.
CM	Completeness.
CR	Correctness.
CVIS	Capsules Visual Inspection System.
DAS	Document Analysis System.
DTREG	Decision Tree Regression.
EC	Expected Characters.
ER	Estimated Region.
FN	False Negative.
FP	False Positive.
GLCD	Gray Level Dependence Matrix.
GLCM	Gray Level Co-occurrence Matrix.
HMM	Hidden Markov Model.
LMS	Least Mean Square.
LDA	Linear Discriminant Analysis.
LoG	Laplacian of a Gaussian.
MET	Minimum Error Thresholding.
MLO	Medio Lateral Oblique.
MLP	Multilayer Perceptron.
MSE	Mean Square Error.
NDT	Non Destructive Testing.
NN	Neural Network.
OCR	Optical Character Recognition.
ON	Ontario.
PC	Personal Computer.
PEIPA	Pilot European Image Processing Archive.
PSNR	Peak Signal to Noise Ration.
QIR	Quadratic Integral Ratio.
RC	Recognized Characters.
ROI	Region of Interest.
SGLD	Spatial Gray Level Dependence.
SN	Sensitivity.
SOFM	Self Organizing Feature Map.
SOM	Self Organizing Map.
SP	Specificity.
TN	True Negative.
TP	True Positive.
USB	Universal Serial Bus.

Chapter 1

Introduction

1.1 Introduction

Image analysis system refers to the extraction of meaningful information from acquired images by means of image processing techniques. In any applications that uses computer vision to automate the respective process, image analysis system plays a great role. The principal stages of an image analysis system is shown in figure 1.1. Each of the steps is an individual research fields and needs intensive study for dealing with low contrast images of complex background for example transparent gelatin capsules in pharmaceutical manufacturing, document images in document recognition and x-ray mammograms in biomedical imaging applications. However, image segmentation is one of the vital steps for the performance of the computer vision-based system as well as pattern recognition in any applications. Success of the vision-based system mostly depends on the quality of the binary image. The complex background, non-uniform illumination and low contrast makes the image segmentation very much challenging. Therefore, image segmentation is an important step and automatic image segmentation makes the process faster, accurate and easier in image analysis system.

Artificial Neural Network (ANN) is one of the important image analysis tools inspired by human perception models. ANN can handles all the challenges in image segmentation for example low contrast, complex background and non-uniform illumination. In this thesis, several methods are developed using ANN and edge-detectors for image segmentation and classification. These are efficient ANN-based thresholding technique and its application in document recognition, efficient automated

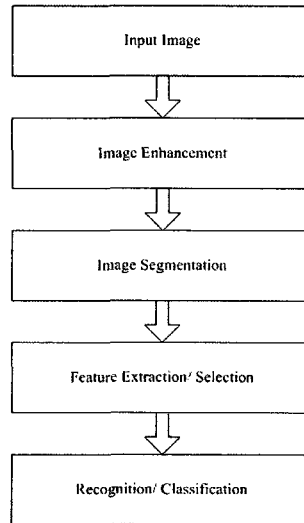


Figure 1.1: Principal stages of image analysis system

seeded region growing for mass extraction in digitized mammogram using Haralick texture features and efficient ANN-based mass classification in biomedical imaging and finally edge-based thresholding and its application in quality inspection of transparent gelatin capsules in pharmaceutical manufacturing applications.

1.2 Image Segmentation

In computer vision, segmentation refers to the process of recognizing homogeneous regions within an image as distinct and belonging to different objects. The segmentation process can be based on finding the maximum homogeneity in gray levels within the regions identified. The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. Because, in computer vision-based system, the digital color image is represented by 24-bit (16- millions levels) or gray scale image (scanned document, X-ray mammograms) by 8-bit (256 levels). The analysis of an image with that many levels might require complicated techniques and higher computational cost. But in real time application an image must be analyzed within a few milliseconds. Therefore, either we have to reduce the image to a more manageable number of gray levels, usually two levels (binary image), and at the same time retain all necessary features of the original image or we have to develop complicated techniques. Thresholding/ multi level thresholding is a way of solving this issue and a binary image can be

obtained by thresholding the image into two levels.

1.3 Quality Process Control of Pharmaceutical Manufactured Products

One of the most important needs, perhaps the most urgent from the point of view of industry productivity and competitiveness, is automatic inspection. Early detection of defects in the production means lower costs and faster feedback on the production line in order to eliminate the causes of defects, overcome physical limitations and subjective judgment of humans. In pharmaceutical industry, zero defect quality capsules are highly required in competitive markets. Therefore, a very cost effective, high throughput and reliable quality inspection method is important in that industry. Two part (body and cap) gelatin capsules are sensitive to several common flaws that make them unacceptable for shipment and marketing. These flaws are incorrect size or color, dents, cracks, holes, bubbles, missing caps and so on.

As consumers are demanding that products be free of defects, 100% inspection and automated inspection have to be applied to achieve the zero-defect and at the same time to maintain the high throughput, all defects must be detected in a single inspection. A computer vision-based quality inspection can be one solution to solve these types of problems. In this thesis we have developed a computer vision-based quality inspection system capable of isolating and inspecting capsules at the rate of 1000 capsules per minute with over 95% accuracy of detecting defects of size 0.2 mm and larger.

1.4 Document Image Analysis

Document image analysis refers to algorithms and techniques that are applied to document images to obtain a computer-readable description from pixel data. Document image contains only raw data that must be further analyzed to process the information. Document image analysis provides fast storage, recall and distribution of documents in work flow processing mostly in business and government applications. In order to provide more flexible retrieval and manipulation capabilities, document analysis system helps to interpret the information represented in the stored images. The most potential application of document image analysis systems are automatic recognition of different forms for example income tax forms, admission forms, postal address and bank cheques reading, license plate recognition and journal, magazines reading.

Giant steps have been taken in the last decade, both in terms of technological supports and in software products to provide computerized DAS. Optical Character Recognition (OCR) contributes to this progress by providing techniques to convert large volumes of texts to readable files automatically. There are so many papers and patents claiming recognition rates as high as 99.99%, this gives the impression that OCR problems seem to have been solved. However, the failure of some real applications shows that performance problems subsist on composite and degraded paper documents with non-uniform background. Non-uniform background is caused by watermarks and complex patterns used in printing security documents. Transforming composite documents with non-uniform background into electronic format in a form suitable for efficient storage, retrieval and interpretation continues to be a challenging problem. Thresholding is a popular tool for image segmentation that tries to identify and extract the object from its background on which it is superimposed. In general, performance of the image thresholding technique depends on the type of document, image illumination, contrast and the complexity of the image background. In this thesis, an exhaustive search was conducted to find the optimal parameter set that was used to develop an efficient ANN-based local thresholding method.

1.5 Biomedical Imaging- Breast Cancer Detection

Breast cancer continues to be a public health problem in the world. It is the second leading cause of death in Canada for women, after lung cancer. In 2009, an estimated 22,700 Canadian women and 170 men were diagnosed with breast cancer and 5,400 women and 50 men died from it. So, 1 in 9 women (11%) is expected to develop breast cancer during her lifetime (by age 90) and 1 in 28 will die from it. Early detection of breast cancer, allowing treatment at an earlier stage, can significantly reduce breast cancer mortality.

Mammography has been one of the most reliable methods for early detection of breast carcinomas. X-ray mammography is currently considered as standard procedure for breast cancer diagnosis. However, retrospective studies have shown that radiologists do not detect all breast cancers that are visible on the mammograms. The estimated sensitivity of radiologists in breast cancer screening is only about 75% [1]. Double reading has been suggested to be an effective approach to improve the sensitivity. But it becomes costly because it requires twice as many radiologists' reading time. This cost will be quite problematic considering the ongoing efforts to reduce costs of the health care system. Cost effectiveness is one of the major requirements for a mass screening program to be successful. Therefore, the main objective of this thesis is to develop a CAD system for breast

cancer diagnosis and detection based on automated segmentation of masses in mammograms. It is expected that the automated methods for classifications provides a tentative diagnosis of individual masses, based on their physical attributes to reduce the number of false positive.

1.6 Thesis Objectives

The main aim of this study is to investigate the potentiality of using ANN in order to develop image processing techniques to extract low contrast objects and classify them to the appropriate classes for the applications in quality process control of pharmaceutical industry manufactured products, efficient ANN-based local thresholding technique in document recognition and computer-aided breast cancer diagnosis and detection from X-ray mammograms in biomedical imaging applications.

In real-time pharmaceutical manufactured products applications, it is difficult to extract an accurate, error free capsule image from the captured image within a very short time especially when the capsules are clear and transparent. Capsule approximation is one of the solutions for this particular problem. From the approximated capsule image, different parameters can be measured to detect the defects like scratches, holes, bubbles, dents, meshed and so on. One of the main objectives of this thesis is to develop a computer vision-based quality inspection system capable of isolating and inspecting capsules at the rate of 1000 capsules per minute with over 95% accuracy of detecting defects of size 0.2 mm and larger.

In document recognition applications, success of OCR mostly depends on the quality of the thresholded image. Non-uniform illumination, low contrast and complex background make it challenging in this application. In this thesis, selection of optimal parameters through exhaustive search for NN-based local thresholding approach for gray scale composite document image with non-uniform background is implemented. To validate the obtained features some non-uniform watermarked document images with known binary document images called base documents were used and a quantitative measure PSNR was used to compare the performance with other methods discussed in this literature.

In biomedical imaging applications, the ultimate goal is to develop CAD system for mass detection and classification in digitized mammograms, which performs mass detection on regions of interest (ROI) followed by the benign-malignant classification on detected masses. In order to detect mass effectively, a sequence of preprocessing steps are proposed to enhance the contrast of the image, remove the noise effects, remove the X-ray label and pectoral muscle and locate the suspicious masses using Haralick texture features generated from the spatial gray level dependence (SGLD) matrix.

The main aim of the CAD system is to increase the effectiveness and efficiency of the diagnosis and classification process in an objective manner to reduce the numbers of false-positive of malignancies.

1.7 Thesis Organization

This thesis is organized as follows: Chapter 2 gives the brief introduction of ANN, different structures, it's importance and it's applications in real life problems. Chapter 3 reviews the different image thresholding techniques, their limitations and strengths. It also describes the statistical and texture features and their use in ANN-based thresholding techniques in real-time and different off-line applications. Chapter 4 introduces the document image analysis system. An exhaustive search is conducted to find the optimal parameters to devise ANN-based efficient thresholding technique, it's validation and the simulation results obtained by the proposed technique is also discussed in this chapter. Chapter 5 introduces the problems in existing quality inspection system (OptiSorter) of gelatin capsules in pharmaceutical manufacturing applications. Current state-of-the-art solutions and the proposed edge-based image segmentation technique to develop image analysis system for this application are also discussed in this chapter. Chapter 6 introduces the importance of computer aided detection of breast cancer. The proposed automatic seed selection method using Haralick texture features for region-growing segmentation technique is developed in this chapter. The extracted mass is then used as an input to the proposed three layers NN and classify the mass into benign or malignant and finally chapter 7 provides concluding remarks and details future work.

Chapter 2

Introduction to Artificial Neural Network

2.1 Introduction

An artificial neural network is an information processing system which has been developed as a generalization of mathematical models of human cognition or neural biology, based on the assumption that:

1. Information processing occurs at many simple computing elements (or neurons).
2. Signals are passed between computing elements over connection links.
3. Each connection link has as an associated weight, which multiplies with the signals transmitted through it.
4. Each computing element applies an activation function to its net input (sum of weighted input signals) to determine its output signals.

ANNs has many names such as connectionist models, parallel distributed processing models, neuro-morphical systems, self-organizing systems and adaptive systems. The true power of ANNs lies in their ability to represent both linear and non-linear systems and in their ability to learn the relationships directly from the data.

ANN's synaptic weights are adjusted or trained so that a particular input leads to a specific desired or target output. Figure 2.1 shows the block diagram for a supervised learning ANN, where the network is adjusted based on comparing neural network output to the desired output until the

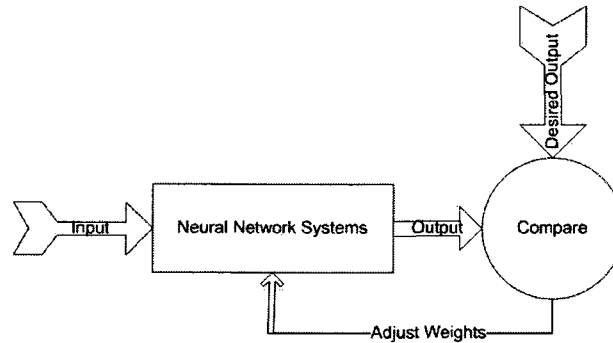


Figure 2.1: Supervised learning of ANN

network output matches the desired output. Once the network is trained it can be used to test new input data using the weights provided from the training session.

2.2 Types of ANNs

ANNs can be classified according to the pattern of connections between the units and the propagation of data. Two main distinctions can be made according to the pattern of connections.

2.2.1 Feed-forward Neural Network

The neurons in this model are grouped in layers, which are connected to the direction of passing signals. Feed-forward ANNs allow signals to travel one way only; from input to output. There is no feedback (loops) i.e. the output of any layer does not affect that same layer. Feed-forward ANNs tend to be straight forward networks that associate inputs with outputs. They are extensively used in pattern recognition. This type of organization is also referred to as bottom-up or top-down. Classical examples of feed-forward neural networks are the Perceptron and Adaline. Figure 2.2 represents the structure of multi-layer feed-forward network.

2.2.2 Feedback Neural Network

Feedback networks can have signals traveling in both directions by introducing loops in the network. Feedback networks are very powerful and can become extremely complicated [2]. Feedback networks are dynamic; their 'state' is changing continuously until they reach an equilibrium point.

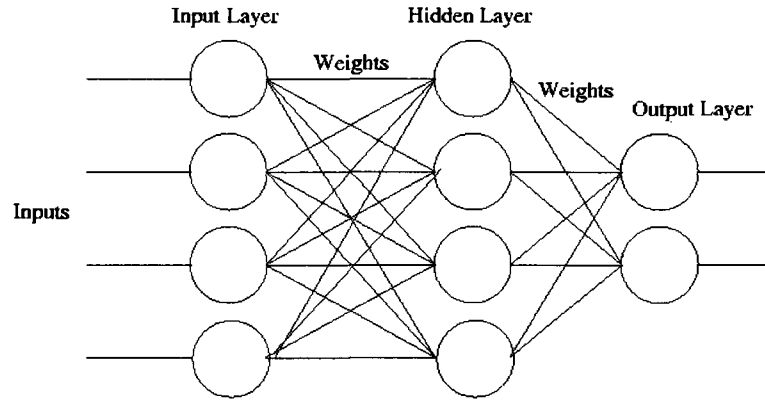


Figure 2.2: Multi-layer feed-forward network model

They remain at the equilibrium point until the input changes and a new equilibrium needs to be found. Feedback architectures are also referred to as interactive or recurrent. Examples of recurrent networks are Anderson networks [3], Kohonen self-organizing maps (SOM) [4], and Hopfield network [5]. Figure 2.3 shows an example of complicated feedforward/ feedback network.

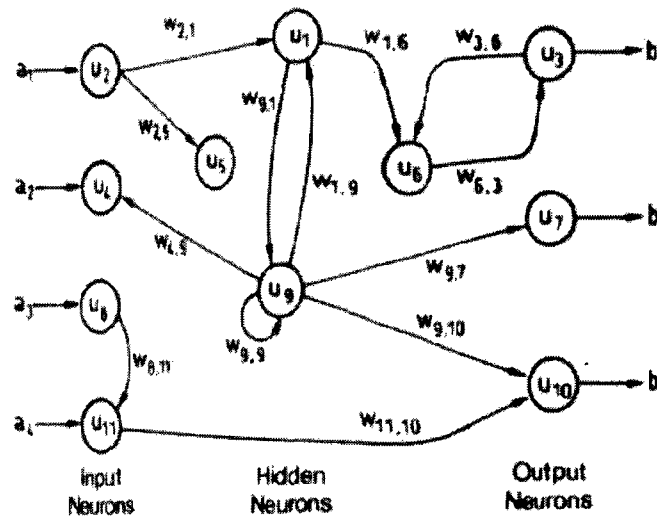


Figure 2.3: Complicated feed-forward/ feedback network

2.3 The Learning Process of ANNs

ANNs learn by adjusting the synaptic weights between neurons, eliminating some synapses and building new ones. All learning methods used for adaptive neural networks can be classified into supervised learning and unsupervised learning. If the learning phase and operation phase are performed in two different times, then it is called off-line learning. If they are performed at the same time then it is called on-line learning. Usually, supervised learning is done off-line and unsupervised learning is performed on-line.

2.3.1 Supervised Learning

In supervised learning, output unit is told what its desired response to input signals ought to be. An ANN of supervised learning, such as MLP uses the target result to guide the performance of the neural parameters. During the learning process global information may be required. Paradigms of supervised learning include error-correction learning, reinforcement learning and stochastic learning. An important issue concerning supervised learning is the problem of error convergence, i.e the minimization of error between the desired and computed unit values. The aim is to determine a set of weights which minimizes the error. One well-known method, which is common to many learning paradigms, is the least mean square (LMS) convergence.

2.3.2 Unsupervised Learning

Unsupervised learning is based upon only local information and uses no external teacher. It is also referred to as self-organization, in the sense that it self-organizes data presented to the network and detects their emergent collective properties. Unsupervised learning of ANNs, for example SOM can be used for clustering the input data and finding features inherent to the problem. Paradigms of unsupervised learning are Hebbian learning and competitive learning.

2.4 Multi-layer Perceptron (MLP)

Multi-Layer perceptron (MLP) is the most common neural network model. It uses supervised training methods to train the network and is structured hierarchically of several perceptrons. MLPs contain normally three layers: the input layer, hidden layers and output layer that is obvious from the figure 2.2. The training of such a network is complicated than single perceptron. That is why when there exists an output error; it is hard to say how much error comes from the different nodes of

input, hidden and output layers and how to adjust the weights according to their contributions [6]. This problem can be solved by finding the effect of all the weights in the network by using the back-propagation algorithm [7] which is the generalization of the least-mean-square (LMS) algorithm. The input nodes, hidden nodes and the output nodes are connected via variable weights using feed-forward connections. The calculated output is compared with the target output. The total mean square error (MSE) shown in equation 2.1, is computed using all training patterns of the calculated and target outputs.

$$MSE = \frac{1}{2} \sum_{j=1}^m \sum_{i=1}^k (T_{ij} - O_{ij})^2 \quad (2.1)$$

where m is the number of examples in the training set, k is the number of output units, T_{ij} is the target output value (either 0.1 or 0.9) of the i – th output unit for the j – th training example, and O_{ij} is the actual real-valued output of the i – th output unit for the j – th training example. The back-propagation algorithm uses an iterative gradient technique to minimize the MSE between the calculated output and the target output. The training process is initialized by setting some small random weights. The training data are repeatedly presented to the neural network and weights are adjusted until the MSE is reduced to an acceptable value. Figure 2.4 shows the training steps of MLP neural network.

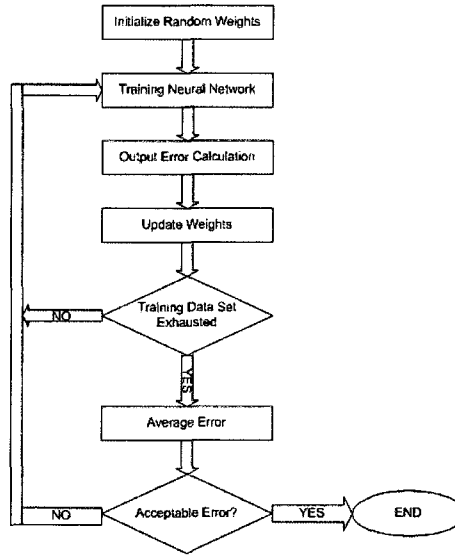


Figure 2.4: Flow chart of the training steps of MLP ANN

2.5 Advantages and Disadvantages of ANNs

A neural network has the ability to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques.

2.5.1 Advantages of ANNs [2]

A trained neural network can be considered as an "expert" in the area of information it has been given to analyze. The expertise of this trained ANN can then be used to handle new situations of interest and answer "what if" questions to that area of interest. Other advantages include:

1. It has the ability to learn how to do tasks based on the data given for training or initial experience.
2. It is able to create its own organization or representation of the information it receives during learning time.
3. ANN computations may be carried out in parallel, and special hardware devices are being designed and manufactured which take advantage of this capability.
4. Partial destruction of a network leads to the corresponding degradation of performance. However, some network capabilities may be retained even with major network damage.
5. Neural models are highly tolerant to noisy data.
6. There is no need to assume an underlying data distribution such as usually is done in statistical modeling.
7. Neural networks are applicable to multivariate non-linear problems.
8. The transformations of the variables are automated in the computational process.

2.5.2 Disadvantages of ANNs [2]

The main disadvantage of neural networks is that they are very slow for large network, especially in the training phase but also in the application phase that makes it unsuitable for on-line processing. Another significant disadvantage of neural networks is that it is very difficult to determine how the net is making its decision. Consequently, it is hard to determine which of the image features being used are important and useful for classification and which are worthless. Other disadvantages include:

1. Minimizing overfitting requires a great deal of computational effort.
2. The relationship between the input and output variables are not developed by engineering judgment so that the model tends to be a black box or input/output table without analytical basis.
3. The sample size has to be large.

4. Global minima is not obtained.

2.6 Application of ANNs

The artificial neural network can be used to infer a function from observations. This is particularly useful in applications where the complexity of the data or task makes the design of such a function by hand impractical. It is applied in broad areas of real world problems in business, industry and medicine [2, 6]. Since neural networks are best at identifying patterns or trends in data, they are well suited for prediction or forecasting like sales forecasting, industrial process control, customer research, data validation, risk management, stock market prediction, bankruptcy prediction, target marketing and dynamic system modeling. More specifically it is also used in recognition of speakers in communications, diagnosis of hepatitis, recovery of telecommunications from faulty software, interpretation of multi-meaning Chinese words, undersea mine detection, texture analysis and three-dimensional object recognition. Other applications include:

1. Signature recognition and verification especially in the bank transactions.
2. Shape recognition of manufactured parts.
3. Handwritten word recognition.
4. Facial recognition.
5. Thresholding or binarization of gray scale or color images.
6. Color recognition.
7. Recognition of items through color/texture.

2.7 Conclusions

ANN is a powerful data modeling tool that provides an alternative and better solutions when expert knowledge is unavailable in full-fledged sense as for example in the application of document recognition and biomedical imaging. The development of the ANN technology from the history shows that it was started from the desire to develop an artificial system that could perform intelligent tasks similar to those performed by the human brain. The human brain is an incredibly impressive information processor, even though it works quite a bit slower than an ordinary computer. Many researchers in artificial intelligence look to the organization of the brain as a model for building intelligent machines like ANN. Among different types of ANNs, feed forward MLP ANN is mostly used

in different applications because of it's simplicity in terms of organizations and learning methods. The use of ANNs in quality inspection of gelatin capsules in pharmaceutical manufactured products, document recognition and biomedical imaging applications are shown in the following chapters.

Chapter 3

Review of Image Thresholding Methods and Statistical Texture Features

3.1 Introduction

To interpret or analyze an image automatically, we must have a way of identifying unambiguously the pixels that corresponds to particular features of interest. The process of identifying these pixels is known as segmentation. It is generally the first stage and one of the critical stages in any attempt to automatic image processing shown in figure 1.1. The role of segmentation is very important in most tasks requiring image analysis. Because the success or failure of the tasks is directly dependent on the success or failure of the image segmentation. Some of the image segmentation techniques are broadly implemented in the applications involving the detection, recognition and measurement of objects in images. These applications include but not limited to quality inspection in industrial manufactured products, OCR, tracking of moving objects, classification of terrains in satellite images, detection and measurement of bone, tissue, cancer in mammograms etc. and in biomedical imaging.

The automatic binarization of gray-level images or the automatic determination of an optimum threshold value is still a difficult and challenging problem in many image processing applications. The difficulty may arise due to various factors, including, poor contrast, non-uniform illumination, composite image with complex background, high noise to signal ratio, complex patterns, and/or variable modalities in the gray-scale histograms. Many thresholding algorithms have been reported in

the literature. The survey of image thresholding can be found in [8, 9, 10]. Some of the thresholding techniques will be briefly described in the following sections.

3.2 Survey Over Thresholding Techniques

To convert the gray scale image into binary image, a threshold is chosen in between 0 (black) and 255 (white). Different binarization methods are discussed in the literatures based on the attributes of the image. Binarization methods can be categorized into six different types. They are as follows:

1. Histogram-based methods
2. Edge-based methods
3. Clustering-based methods
4. Entropy-based methods
5. Local adaptive thresholding methods
6. Neural network-based methods.

3.2.1 Histogram-based Methods

Histogram-based methods are very simple and efficient compared to other methods because of its low computation cost and zero memory storage requirements. In this approach, a histogram is computed from all of the pixels in the image, and the peaks, valleys and curvature of the smoothed histogram of the image are analyzed to locate the clusters in the image. Color or intensity can be used as the measure. A refinement of this technique is to recursively apply the histogram-seeking method to clusters in the image in order to divide them into smaller clusters. This is repeated with smaller and smaller clusters until no more clusters are formed [11, 12]. One disadvantage of the histogram-seeking method is that it may be difficult to identify significant peaks and valleys in the image. However, in most applications the threshold is determined from the peak and valley detection. For example, Sezan used peak [13], Sahasrabudhe [14] and Ohio [15] used valley to determine the threshold. Rosenfeld [16] used the analysis of concavities of the histogram that is called convex hull and Weszka [17] used enhanced concavity by modifying its histogram. When the convex hull of the histogram is calculated, the deepest concavity points become candidates for a threshold. Rosin [18] developed an algorithm that fits a straight line from the peak of the histogram to the last non zero

bin. The point of the maximum deviation between the line and the histogram is selected as the threshold value which is the deepest valley point.

Tsai [19] developed a moment-preserving technique for threshold selection where first three moments of both input and output images are preserved. The moment-preserving technique can be thought of as an information transformation method which groups the pixels of an image into classes. The variables in the so-called moment-preserving equations are determined iteratively by a recurrent neural network and a connectionist neural network which work cooperatively. Both of the networks are designed in such a way that the sum of square errors between the moments of the input image and those of the output version is minimized. Initially the histogram is smoothed via Gaussian function, and the resulting histogram is investigated for the presence of both valleys and sharp curvature points. The curvature analysis becomes effective when the bimodality of the histogram is lost due to the excessive overlapping of class histograms.

The Quadratic Integral Ratio (QIR) [20] is a global two stage histogram-based thresholding approach. Initially the image histogram is divided into three classes, object, background and fuzzy class. Fuzzy class is in between the object and background where it is hard to tell in which class the pixel belongs. At second stage, final threshold is determined from the fuzzy region. QIR method performs well on constant homogeneous background of the image possessing bimodal histogram. Leedham et. al. [21] concluded that the QIR method is more accurate in separating object from the background than other multi-stage global thresholding methods. Other techniques include shape of the histogram [22], and histogram modification via partial differential equations [23].

The histogram based methods described above are mostly global thresholding method that uses single threshold for the entire image. These methods shows very good performance provided that the histogram of the image is clearly bimodal which is very unlikely for real world images with noisy, non-uniform background. In that case, local thresholding [24] is one of the ways to solve the bimodal issue where the threshold is changing over the image. Local thresholding methods divide the image into multiple non-overlap regions so that the pixels are considered homogeneous in each region. This way it tackles the non-uniformity of the image background. But in these methods it is difficult to determine the correct window size yielding incorrect threshold value based on the contrast and illumination of the background. It has another drawback called blocking effect that happens when the difference of the threshold level between two adjacent blocks are very large yielding discontinuity. Some of the developed local thresholding methods are multi-level thresholding methods [25, 26], adaptive thresholding methods [27, 28] and iterative thresholding methods [29, 30].

3.2.2 Edge-based Methods

Edges are defined as the transition of gray values from object to background and the pixels around edge are the mixture of the object and background. Edge-based methods are based on the hypothesis that gray level of the pixels around the boundary of the object and background changes abruptly no matter how complex the background is. Edge-based method is a two stage process that is shown below:

1. Edge detection.
2. Object and background pixels localization.

The discontinuity of an edge is detected by the first and second order derivatives. In image processing, the first order derivative is called the gradient. The gradient of 2D function $f(x, y)$ is defined as a vector shown in equation 3.1. The magnitude of this vector is calculated using the equation 3.2 and the gradient direction is calculated using the equation 3.3. G_x and G_y are the convolutions of the original image and the mask in horizontal and vertical direction respectively. There are six edge detectors that are mostly used in different applications. They are Sobel, Prewit, Roberts, Laplacian of a Gaussian (LoG), Zero crossing and Canny [31]. Sobel finds the edges using the horizontal and vertical mask shown in table 3.1 and 3.2. The horizontal and vertical mask for Prewit edge detector is shown in table 3.3 and 3.4. Although Prewit or Sobel filters are larger they are less sensitive to noise. Prewit is easy to implement but Sobel has slightly superior noise-suppression characteristics. Roberts edge detector is fast since the mask is small but it is also subject to interference by noise. If edges are not very sharp the edge detector will tend not to detect the edge. Robert's mask are shown in tables 3.5 and 3.6 respectively. LoG and zero crossing detects the edges by looking for zero crossing after filtering $f(x, y)$ with a Gaussian filter and user-specified filter respectively. Canny edge detector looks for local maxima of $f(x, y)$ and $f(x, y)$ is calculated using the derivative of a Gaussian filter. Then strong and weak edges are detected using two threshold values and finally the weak edges connected to the strong edges are included in the output.

$$\nabla \mathbf{f} = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \quad (3.1)$$

$$\nabla f = \text{mag}(\nabla \mathbf{f}) = \sqrt{G_x^2 + G_y^2} \quad (3.2)$$

$$\Theta = \tan^{-1} \left(\frac{G_y}{G_x} \right) \quad (3.3)$$

Table 3.1: Horizontal mask of Sobel Edge Detector

-1	-2	-1
0	0	0
1	2	1

Table 3.2: Vertical mask of Sobel Edge Detector

-1	0	1
-2	0	2
-1	0	1

Table 3.3: Horizontal mask of Prewitt Edge Detector

-1	-1	-1
0	0	0
1	1	1

Table 3.4: Vertical mask of Prewitt Edge Detector

-1	0	1
-1	0	1
-1	0	1

Table 3.5: Roberts' Edge Detector (-45°)

0	-1
1	0

Table 3.6: Roberts' Edge Detector ($+45^\circ$)

-1	0
0	1

Once the edge strength is measured, the next stage is to apply a threshold, to decide whether edges are present or not at an image point. The lower the threshold, the more edges will be detected, and the result will be increasingly susceptible to noise, and also to picking out irrelevant features from the image. Conversely a high threshold may miss subtle edges, or result in fragmented edges. Some of the developed edge based methods for different applications are described below:

Huang [32] developed edge based binarization technique for the application of document recognition. In this method once edges are detected, some pixels around edges are selected to represent object and background. Since most of the disturbing background information is ignored after the edge detection, the selected pixels around edges form a bimodal histogram from where the threshold value can be selected easily.

Prager [33] proposed a set of algorithms used to perform segmentation of natural scenes through boundary analysis. The goal of the algorithm is to locate the boundaries of an object correctly in a scene. In this method differentiation is done to find the edge-strength at each point in the image. Then suppression is done to remove multiple edges formed by spatial differentiation of boundaries. Then the edges are joined into line segments and the features length, contrast, frequency, mean, variance and location of each line segment are computed. Finally, post-processing is done to remove unwanted line segments and to build confidence for each of the remaining segments.

Perkins [34] uses an edge based technique for image segmentation. Edge based segmentation has not been very successful because of small gaps that allow merging of dissimilar regions. In order to avoid these problems, Perkins proposes an expansion contraction technique in which edge regions are expanded to close gaps and then contracted after the separate regions have been labeled. The size of expansion is controlled such that small regions are not engulfed by this process. The process involves the use of Sobel edge detector for producing edge strengths and directions at every point. The edges are thinned and the result is automatically thresholded leaving only ridges. The ridges separate regions of different intensity but there may be small gaps. This method was used for the application of landscape pictures and for pictures of electronic circuits.

Yalcyn et. al. [35] uses an edge based approach to real time automated visual inspection for post manufacturing industrial parts quality inspection. In this system, the image is selectively processed and the salient contour segments are implicitly identified as a sequence of fixation points. This not only reduces the required computation drastically, but also enables merging of the segments and interpolation of the contour in a natural manner.

3.2.3 Clustering-based Methods

The gray-level samples are clustered into two classes as background and object, or alternatively they are modeled as a mixture of two Gaussians in the clustering-based methods. The Otsu method [37] is the most cited clustering-based thresholding method. In this method the weighted sum of within-class variances that turns out to be the maximization of between-class variances. It operates directly on the gray level histogram $P(i)$. This method provides satisfactory results when the number of pixels in background and objects are similar and fails in images with complex background or degraded images.

The weighted within-class variance is defined by:

$$\sigma_w^2(t) = q_1(t)\sigma_1^2(t) + q_2(t)\sigma_2^2(t) \quad (3.4)$$

where the class probabilities are estimated as:

$$q_1(t) = \sum_{i=1}^t P(i) \quad (3.5)$$

$$q_2(t) = \sum_{i=t+1}^n P(i)$$

and the class means are given by

$$\mu_1(t) = \sum_{i=1}^t \frac{iP(i)}{q_1(t)} \quad (3.6)$$

$$\mu_2(t) = \sum_{i=t+1}^n \frac{iP(i)}{q_2(t)}$$

Finally the individual class variances are defined by:

$$\sigma_1^2(t) = \sum_{i=1}^t [i - \mu_1(t)]^2 \frac{P(i)}{q_1(t)} \quad (3.7)$$

$$\sigma_2^2(t) = \sum_{i=t+1}^n [i - \mu_2(t)]^2 \frac{P(i)}{q_2(t)}$$

Otsu method is implemented in iterative approach. It runs through all grey levels and pick the value that minimizes weighted within-class variance $\sigma_w^2(t)$.

Kittler [38, 39] developed a computationally efficient thresholding approach called minimum error thresholding (MET) for gray scale image binarization under the assumption that the object and background pixel gray level values being normally distributed. This method optimizes the average pixel classification error rate directly, using an iterative algorithm.

In the MET technique to find the optimal threshold τ it uses the simpler technique. Suppose that the histogram of the gray scale image is $h(g)$ assuming gray level values, g , from the interval $[0..n]$. The trial threshold is set at arbitrary level T . It models the two resulting pixel populations with parameters $\mu_i(T)$, $\sigma_i(T)$ and an *a priori* probability $P_i(T)$ given, respectively, as

$$P_i(T) = \sum_{g=a}^b h(g) \quad (3.8)$$

$$\mu_i(T) = \left[\sum_{g=a}^b h(g)g \right] / P_i(T) \quad (3.9)$$

and

$$\sigma_i^2(T) = \left[\sum_{g=a}^b \{g - \mu_i(T)\}^2 h(g) \right] / P_i(T) \quad (3.10)$$

where

$$a = \begin{cases} 0 & i = 1 \\ T + 1 & i = 2 \end{cases} \quad (3.11)$$

and

$$b = \begin{cases} T & i = 1 \\ n & i = 2 \end{cases} \quad (3.12)$$

If an image has n grey levels, this method tries in succession n different levels as a threshold value. For each grey level T , a fitting criterion $J(T)$ is calculated, which is defined by

$$J(T) = 1 + 2 [P_1(T) \log \sigma_1(T) + P_2(T) \log \sigma_2(T)] \\ - 2 [P_1(T) \log P_1(T) + P_2(T) \log P_2(T)] \quad (3.13)$$

The criterion function $J(T)$ in equation 3.13 can be computed easily and finding its minima is a relatively simple task, as the function is smooth.

Chow and Nakagawa [40, 41] divides the image into non-overlapped regions. The histogram of each region is approximated by a mixture of two Gaussian distributions to test the bimodality. Thresholds are calculated based on the means and standard deviations of the bimodal mixture distribution for each region. Thresholds are first smoothed and then interpolated to form a threshold

surface yielding a single threshold value for each image pixel. Finally the image is binarized using the threshold surface.

Other iterative clustering approaches are found Ridler [29], Yanni [26], Lloyd [42] and Cho [43]. Reddi et. al. [44] developed a fast search scheme based on the criteria proposed by Otsu [37] that maximizes inter-class variances between object and background. Sid-Ahmed [45] proved mathematically that the problem could be reduced to simple iterative algorithms conjectured by Ridler [29]. Fuzzy clustering [46] is another thresholding technique where fuzzy clustering memberships are assigned to pixels depending on their differences from the two class means.

3.2.4 Entropy-based Methods

Entropy is a statistical measure of randomness that can be used to characterize the texture of the input image. Entropy-based methods exploit the entropy of the distribution of the gray levels in an image. Maximum entropy of the binary image is interpreted as an indication of maximum information transfer. Kapur et.al. [36] maximize the sum of entropy so that when the sum of the two class entropies reaches its maximum, the image is said to be optimally thresholded.

Abutaleb [47] used spatial entropy information to generate the optimal threshold followed by two-dimensional thresholding to classify the pixels. Other entropy-based thresholding methods are entropic thresholding [48, 49], cross entropic thresholding [50, 51], fuzzy entropic thresholding [52] and fusion of three different entropies [53].

3.2.5 Local Adaptive Thresholding

Local statistics like range, variance, or surface-fitting parameters of the neighboring pixels are used to calculate the threshold value of each pixel in local adaptive thresholding algorithm. Niblack [54] calculates the pixel-wise threshold value by sliding a window over the image. The threshold value of center pixel (x, y) of a window is calculated from the mean μ and standard deviation σ using the following linear equation 3.14.

$$T(x, y) = \mu(x, y) + k\sigma(x, y) \quad (3.14)$$

where k is a constant which is highly tunable to extract objects from the background. Also, the size of the window determines the quality of the binary image. Window size should be small enough to serve local details and at the same time large enough to suppress noise. The value of k is in between 0 and 1 that determines how much of the object edge is retained as a part of the given

object.

Sauvola [55] is an improved modification of Niblack [54] method. It shows good performance on documents where the background contains light texture, big variations and non-uniform illumination. The threshold is calculated using the following equation 3.15.

$$T(x, y) = \mu(x, y) * \left(1 + k \left(\frac{\sigma(x, y)}{R} - 1\right)\right) \quad (3.15)$$

where R is an empirical constant. It is suggested by the author that the value of k=0.5 and R=128 to be used in stained and badly illuminated documents. A surface fitted to the gray-level landscape is used as a local threshold, as in Yanowitz and Bruckstein [56] and Shen and Ip [57]. Local contrast is exploited by White et. al [58] where it compares the gray value of the pixel with the mean gray level in some neighborhood. If the pixel value is significantly smaller than the mean, it is classified as object otherwise background.

3.2.6 Artificial Neural Network-based Methods

Very few researchers have investigated the use of neural network in thresholding of gray level images with the recent development of artificial intelligence. Artificial neural network (ANN) based image thresholding method calculates the threshold value based on some local statistics like mean, standard deviation, smoothness, entropy and so on. Yan [59] introduced a technique to discriminate the pixels into background and object according to the characteristics around every pixel. The major drawback of such a method is that binarization is based on the feature around every pixel in an image, therefore the computational cost is much higher than others which makes it unsuitable for online applications.

Koker and Sari [60] used ANN to automatically select a global threshold value for an industrial vision system based on the histogram of the image. The image is preprocessed by using a median filter to remove the effects of noise. In this method the histogram of the supervised training data (256 levels) is used as input and global threshold value that is determined by visually inspecting the histogram of the training image is used as target value. Using this procedure the training data is prepared in different illumination environment for this particular application. In training phase, the produced output is compared to the target output, calculates the error, adjust the synaptic weights of the ANN until an acceptable weight is achieved to minimize the error between produced and target outputs. This method works for this specific application and it fails for composite image with complex background.

Khashman [61] used a supervised neural network for document image binarization where local threshold values are used to train the network in order to obtain the optimum global threshold values.

86% recognition rate is obtained based on visual inspection for degraded non-uniform historical document images.

Papamarkos [62] used Self Organizing Feature Map (SOFM) to define two bi-level classes. Then fuzzy membership functions are obtained from the contents of these classes to use with the fuzzy C-mean algorithm for reducing the character blurring effect. Laplacian and mean values from 3x3 window is used as input to the network. This method works well for blurred documents but not as well for images with complex background.

Alginahi et. al. [7, 63] developed ANN-based technique for thresholding composite digitized documents with non-uniform and complex background. It was used in the application of segmenting bank cheques from complex background for application in OCR. This method uses 8 statistical and textural features of an image for each pixel at a window size 5x5. A modified ANN-based thresholding technique was developed by Alginahi [63] very recently where the author claimed only 5 out of 8 features at window size 3x3 provides the best results. But the author did not present any justification in support of the claim. The statistical and texture features and associated NN-based thresholding method is described in the sections 3.4 and 3.5 respectively.

3.3 Performance Comparison of the Thresholding Techniques

Many image thresholding techniques have been developed and reviewed in the literature and performance of each technique depends on the type of the image and it's complexity of the image background.

Otsu [37] is the most referenced thresholding technique among all global thresholding techniques.

Sahoo et. al. [53] compared the performance of more than 20 global thresholding methods. The comparison shows that the Otsu's [37] method was one of the better threshold selection methods. However, Otsu's [37] method uses an exhaustive search method to maximize the between-class variance. As the number of classes increases, Otsu's method takes too much time to be practical for multilevel threshold selection.

Rosin [64] evaluated the performance of global thresholding methods for change detection and found that Otsu [37] and Riddler [29] shows poor performance compared to other methods.

Fisher [65] compared 15 global thresholding methods and concluded that the Otsu [37] is highly preferred in document image processing.

Sezgin and Sankur [8] conducted an exhaustive survey on 40 bi-level thresholding methods and categorized them into six different groups according to the information they are exploiting. These

methods are compared in the context of nondestructive testing applications as well as for document images. In both cases, the Minimum Error Thresholding (MET) technique proposed by Kittler and Illingworth [38] was ranked as the top one.

Trier and Jain [66] conducted a survey and evaluated the performance of 11 established local adaptive thresholding methods. These methods were compared using a criterion based on the ability of an OCR module to recognize handwritten numerals from hydro-graphical images. In this survey, Niblack [54] was found to be the best one among 11 local adaptive thresholding methods.

3.4 Statistical and Texture Features

Statistical features are extracted for each centered pixel in a window of size $(2n + 1) \times (2n + 1)$, where $n \geq 1$. This window can be randomly placed at different locations by the user in the area of the image in which features need to be extracted. The extracted features are then used in neural classifier to train it for the recognition of a particular area in images of similar nature. The ANN-based thresholding technique takes advantage of the image textural characteristics by considering the statistical descriptors in a neighborhood of pixels. The statistical textural features are useful in characterizing the set of neighborhood values of pixels, which are related to its moments. These features are pixel value, mean, standard deviation, smoothness, entropy, skewness, kurtosis and uniformity. These are adapted from [6, 7, 63, 67].

3.4.1 Pixel Value¹

The center pixel, $p(i, j)$ in a window of size $(2n + 1) \times (2n + 1)$, where $n \geq 1$, was taken as the first feature.

3.4.2 Mean²

The mean, μ_{ij} of the pixel values in the defined window, estimates the value around the pixel in which central clustering occurs.

$$\mu_{ij} = \frac{1}{(2n + 1)^2} \sum_{x=i-n}^{i+n} \sum_{y=j-n}^{j+n} p(x, y) \quad (3.16)$$

Where $p(x, y)$ is the pixel value at point (x, y) .

3.4.3 Standard Deviation³

The standard deviation, σ_{ij} is the estimate of the mean square deviation of grey pixel value $p(x, y)$ from its mean value μ_{ij} . Standard deviation describes the dispersion within a local region. The standard deviation is defined as:

$$\sigma_{ij} = \frac{1}{(2n+1)} \sqrt{\sum_{x=i-n}^{i+n} \sum_{y=j-n}^{j+n} (p(x, y) - \mu_{ij})^2} \quad (3.17)$$

3.4.4 Smoothness⁴

Relative smoothness, R_{ij} is a measure of grey level contrast that can be used to establish descriptors of relative smoothness.

$$R_{ij} = 1 - \frac{1}{1 + \sigma_{ij}^2} \quad (3.18)$$

Where σ_{ij} is the standard deviation around the center pixel (i, j) .

3.4.5 Entropy⁵

Entropy, h_{ij} can also be used to describe the distribution variation in a region. Overall entropy of neighborhood pixels in the window can be calculated as:

$$h_{ij} = - \sum_{k=0}^{L-1} Pr_k (\log_2 Pr_k) \quad (3.19)$$

Where Pr_k is the probability of the k-th grey level, which can be calculated as $Z_k / (2n+1)^2$, Z_k is the total number of pixels with the k-th grey level and L is the total number of grey levels in the window.

3.4.6 Skewness⁶

Skewness, S_{ij} characterizes the degree of asymmetry of a pixel distribution in the specified window around its mean. Skewness is a pure number that characterizes only the shape of the distribution.

$$S_{ij} = \frac{1}{(2n+1)^2} \sum_{x=i-n}^{i+n} \sum_{y=j-n}^{j+n} \left[\frac{p(x, y) - \mu_{ij}}{\sigma_{ij}} \right]^3 \quad (3.20)$$

Where $p(x, y)$ is the pixel value at point (x, y) , μ_{ij} and σ_{ij} are the mean and standard deviation around the center pixel (i, j) respectively.

3.4.7 Kurtosis⁷

Kurtosis, K_{ij} measures the peakness or flatness of a distribution relative to a normal distribution. The conventional definition of kurtosis is:

$$K_{ij} = \left\{ \frac{1}{(2n+1)^2} \sum_{x=i-n}^{i+n} \sum_{y=j-n}^{j+n} \left[\frac{p(x,y) - \mu_{ij}}{\sigma_{ij}} \right]^4 \right\} - 3 \quad (3.21)$$

Where $p(x,y)$ is the pixel value at point (x,y) , μ_{ij} and σ_{ij} are the mean and standard deviation around the center pixel (i,j) respectively. The -3 term makes the value zero for a normal distribution.

3.4.8 Uniformity⁸

Uniformity, U_{ij} is a texture measure based on histogram and is defined as:

$$U_{ij} = \sum_{k=0}^{L-1} Pr_k^2 \quad (3.22)$$

Where Pr_k is the probability of the k -th grey level. Because the Pr_k have values in the range $[0, 1]$ and their sum equals 1, U is maximum in which all grey levels are equal, and decreases from there. Before computing any of the descriptive texture features above, the pixel values of the image were normalized by dividing each pixel by 255 in order to achieve computational consistency.

3.5 ANN-based Thresholding Using Statistical Texture Features

The statistical and textural features described in the section 3.4 are useful in characterizing the set of neighborhood values of pixels. Alginahi [7, 63, 67] utilized these features value in preparing data for training a Multi-Layer Perceptron (MLP) ANN which are obtained from window of size $(2n+1) \times (2n+1)$, taken from various parts of one image and repeated over many images. Figure 3.1 shows a sample screen capture of the training data preparation. The first step to prepare data for training is to load an image then a point or pixel in the image is selected. This is supervised training data preparation method. Therefore, the user knows about the selected pixel whether it is background or object. Accordingly, the user clicks on the object or background buttons in order to calculate the feature vector of that point. The process is repeated for different random points in the same image and different images with complex backgrounds to get a wide variety of features. The 8 features and their corresponding target value (for background=1 and object=0) are stored

Table 3.7: Neural Network Structure used by Alginahi [7]

Input Units	8
Hidden Units	13
Output Unit	1
Weights	117
Image	10
Training Samples	1000

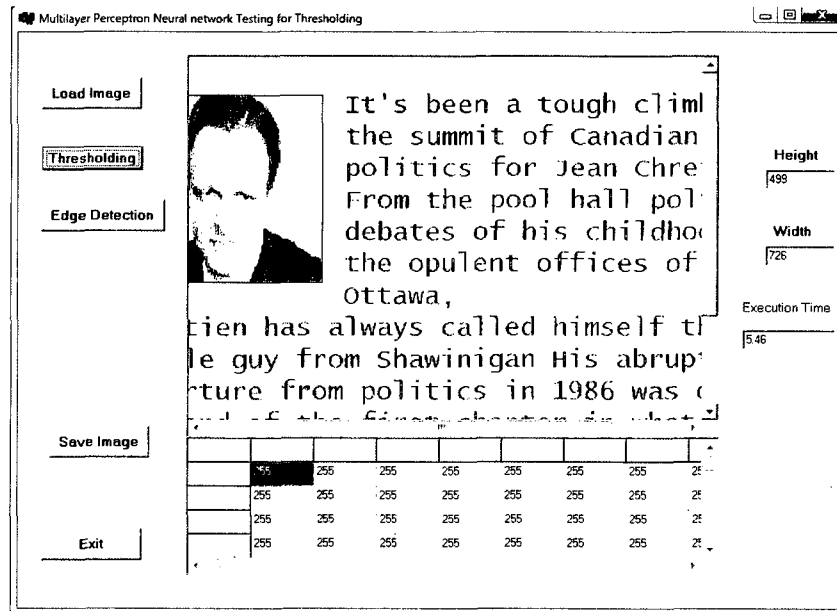


Figure 3.2: Sample Binary Image using NN-based Thresholding

termed as global and local thresholding methods. Among all of these methods surveyed in the literature, it was found that the clustering based method proposed by Otsu [37] is the most popular global thresholding methods. In one survey, Kittler [38] is also found one of the best methods where 40 bi-level methods were tested on 40 nondestructive testing images (NDT) and 40 document images. Trier and Jain [66] found that Niblack [54] method is the best method among 11 well-established local adaptive thresholding methods.

ANN-based thresholding technique is rarely used in online application because of it's high computational cost. However, in document recognition application Alginahi [7, 63, 67] utilized 8 statistical and texture features to develop ANN-based thresholding technique. The correct recognition rate 98% is achieved using those 8 features at window size 5x5. In another research [63] the same author claimed 5 features at window size 3x3 provide the recognition rate 98.3%. But the author failed to provide any justification in support of this claim. In the next chapters we will use ANN-based thresholding technique to investigate it's potentiality in different applications and at the same time we will do exhaustive search to obtain the optimal features subset at optimal window that provides the highest recognition rate.

Chapter 4

Efficient ANN-based Thresholding in Document Recognition Application

4.1 Introduction

For human transactions writing is considered to be more natural way than speech. Printed documents, such as newspapers, magazines and books, and in handwritten matter, such as found in notebooks and personal letters, they all contain the language in written form. Therefore, given the importance of written language in human transactions, its automatic recognition has practical significance. Document Analysis System (DAS), a small division of digital image analysis converts a paper-based document into computerized form. The principal stages of DAS is shown in figure 4.1. Transforming documents with poor contrast and complex background into binary image suitable for efficient storage, retrieval and interpretation continues to be a challenging problem [7, 63, 67]. Sample document image in different complexities are shown in figures 4.2, 4.3 and 4.4.

4.2 ANN-based Thresholding- Observations and Criticism

Image thresholding is one of the very important steps in DAS shown in figure 4.1. Because, it is vital for the performance of DAS and OCR performance can be degraded if the characters in binarized image are broken, blurred or overlapped. In general, performance of the image thresholding

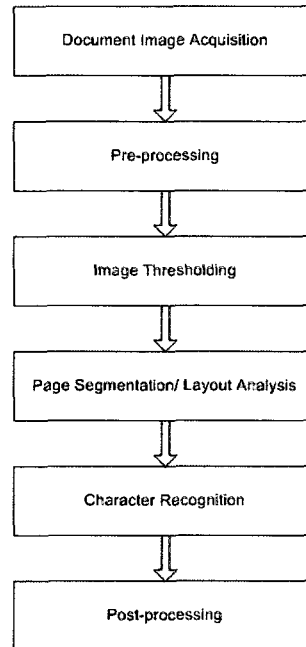


Figure 4.1: Principal Stages of Document Analysis System

The Underground Railroad

In the days before and during the American Civil War, Ontario served as the final stop on the underground railroad, a network of secret routes and safe houses that allowed enslaved African-Americans to escape to freedom in Canada.

Walk in the footsteps of history along the African Canadian Heritage Route from Windsor, where you can visit John Freeman Wall's 1846 log cabin, that served as a terminal on the Underground Railroad. For a further window into the past, walk among the artifacts and images at the Amherstburg's North American Black Historical Museum. Stroll the streets of North Buxton, Canada's first Black settlement - home to many historic buildings and a museum that recounts the area's proud story of early growth and self-sufficiency.

Figure 4.2: Sample Document Image- (poor contrast)

Health is defined as a state of complete physical, social and mental well-being, and not merely the absence of disease or infirmity. Within the context of health promotion, health has been considered less as an abstract state and more as a means to an end which can be expressed in functional terms as a resource which permits people to lead an individually, socially and economically productive life. Health is a resource for everyday life, not the object of living. It is a positive concept emphasizing social and personal resources as well as physical capabilities.

Figure 4.3: Sample Document Image- (non-uniform illumination)

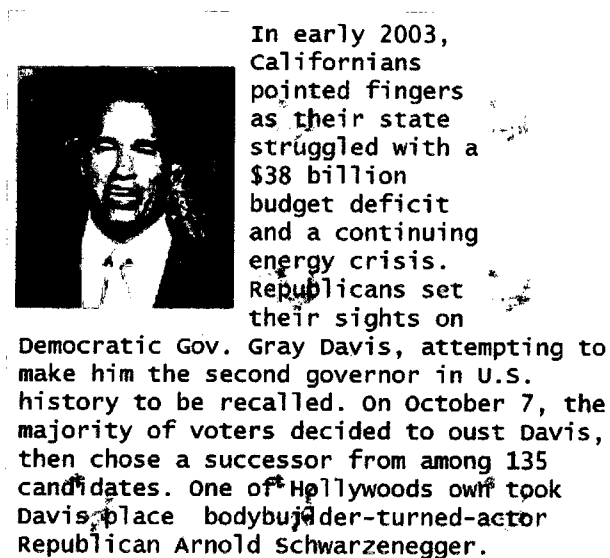


Figure 4.4: Sample Document Image- (composite with complex background)

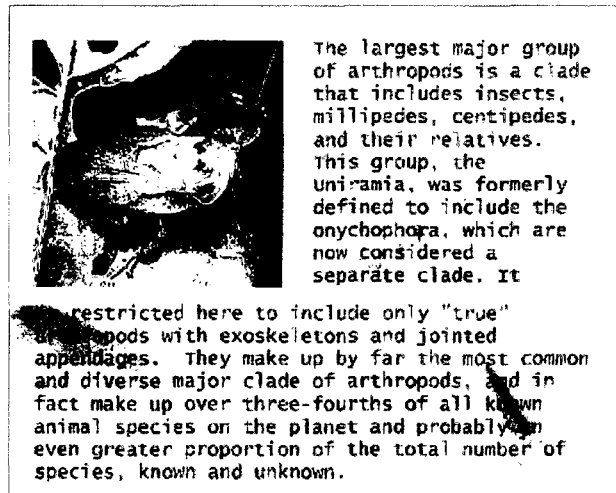


Figure 4.5: Sample Document Image- (Insect)

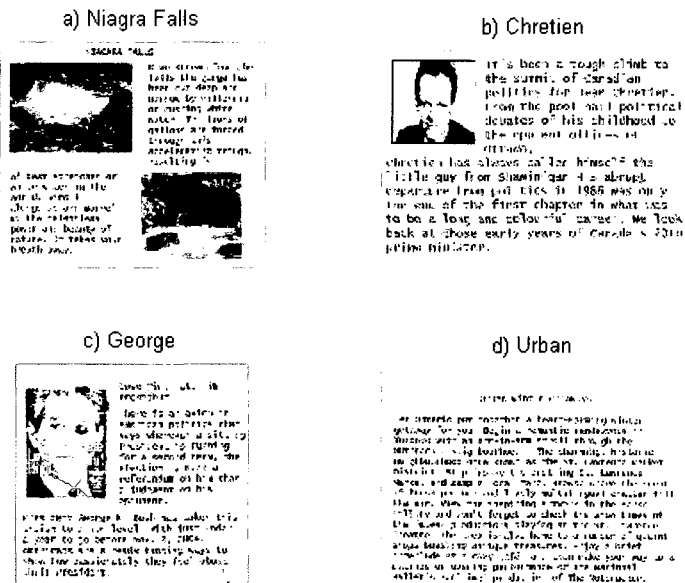


Figure 4.6: More Sample Document Images 1

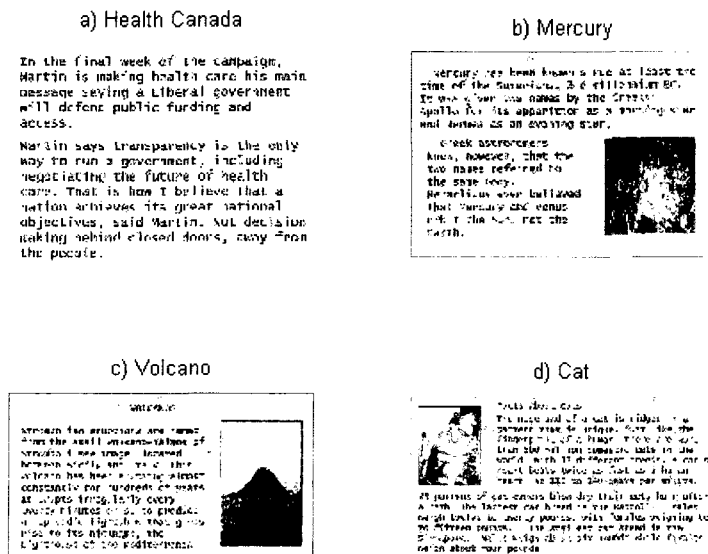


Figure 4.7: More Sample Document Images 2

technique depends on the type of document, image illumination, contrast and the complexity of the image background. There are many thresholding schemes published in the literature in chapter 3 and selecting an appropriate one can be a difficult task. ANN-based local thresholding method [67, 63] is considered to be one of the best solutions in document recognition application in terms of recognition rate that is described in section 3.5 compare to other well established thresholding methods like Otsu's global thresholding [37], Niblack [54], HMM [32] and Kittler [38].

Alginahi et. al. [7, 67] developed ANN-based technique for converting composite digitized documents with complex background into binary form. This method uses 8 statistical and textural features of an image for each pixel that are calculated in a pixel's neighborhood using the window size 5x5. A modified NN-based thresholding technique was developed by Alginahi [63] where only 5 out of 8 parameters at window size 3x3 provides the best results. But the author did not present any justification in support of the claim.

Furthermore, increase in features results in more computational cost and larger window size also increases the computational cost hence slower the feature extraction process. In this thesis, we are questioning the validity of the number of features and window size used in the method [67, 63], whether these 8 or 5 features with the window size of 5x5 or 3x3 are optimal in terms of yielding the

best recognition rate and providing the best peak-signal-to-noise ratio (PSNR). Therefore, the main objective of this thesis is to develop an efficient thresholding technique based on the work reported in [67, 63]. This work therefore presents a method to select the optimal feature set for a neural network based thresholding technique. To fulfill the objectives, an exhaustive search is conducted to find and validate the optimal window size and the optimal feature combination among 8 features that provide the best recognition rate and PSNR when it is compared to other similar techniques available in the literature.

4.3 Proposed Method

According to the above observations and analysis described in section 4.2, we here propose an efficient ANN-based thresholding method based on the optimal parameter set. The proposed method is also called local pixel characteristics based method. The main focus of this method is to extract characters from the image with non-uniform illumination, low contrast and composite with complex background. The proposed method is developed in two phases. At first phase we will present a method to select the optimal feature set for a neural network based local thresholding technique and in the second phase an efficient thresholding technique based on the optimal parameters set obtained in the first phase will be proposed.

4.3.1 Phase I-Optimal Parameters Selection

A pixel can be characterized by the local statistical and textural features extracted from that pixel's neighborhood. Alginahi [7, 63, 67] utilized this concept and extracted 8 features for each pixel from pixel's neighborhood preparing data for training a MLP ANN. To select the optimal parameters, at first these 8 features are labeled as 1(section 3.4.1), 2 (eq. 3.16), 3 (eq. 3.17), 4 (eq. 3.18), 5 (eq. 3.19), 6 (eq. 3.20), 7 (eq. 3.21) and 8 (eq. 3.22) to refer them using a number in the following sections. The next part is to calculate how many different subsets are possible without repetition using those 8 features. Total 255 subsets are possible without repetition from this set of 8 features that are determined using the well known mathematical expression shown in equation 4.1 and number of subsets are shown in figure 4.8. In this method, we are considering 3x3, 5x5, 7x7 and 9x9 window size for each subset to extract features of each pixel. Therefore, for each 255 subsets and 4 window size we have total $255 \times 4 = 1020$ cells that form a grid. The optimal feature subset and window size will be obtained from this grid shown in figure 4.9.

$${}^nC_r = \frac{n!}{r!(n-r)!} \quad (4.1)$$

$$\begin{array}{ll} {}^8C_8 = 1 & {}^8C_4 = 70 \\ {}^8C_7 = 8 & {}^8C_3 = 56 \\ {}^8C_6 = 28 & {}^8C_2 = 28 \\ {}^8C_5 = 56 & {}^8C_1 = 8 \end{array} \quad {}^8C_7 = 8 = \left\{ \begin{array}{l} 1,2,3,4,5,6,7 \\ 1,2,3,4,5,6,8 \\ 1,2,3,4,5,7,8 \\ 1,2,3,4,6,7,8 \\ 1,2,3,5,6,7,8 \\ 1,2,4,5,6,7,8 \\ 1,3,4,5,6,7,8 \\ 2,3,4,5,6,7,8 \end{array} \right\}$$

Figure 4.8: No. of features subset for each combination

For a specific window size, a subset is picked among 255 subsets. Features are extracted according to the subset from the main training dataset for that specific window size that generates a feature vector. The feature vector is used as an input to train the ANN [7] and the corresponding weights are calculated that are stored in a weight vector. The weight vector is fitted to the ANN and is not changing during classification phase. Then the features are extracted from every pixel's neighborhood of the image being tested using the same window size. Extracted features are then inputted to the ANN classifier to classify the pixels belonging to object or background. This process is repeated for all the pixels of the image. Once all pixels are tested and assigned either 0 or 1 it forms a binary image.

The binary image is then passed to the commercial OCR for example ABBYY (version 7.0) to obtain the recognized characters. The recognized characters (RC) are then compared to the expected characters (EC). The expected characters database can be generated by directly feeding the document to the commercial OCR and correcting any misclassified characters from the original image. The error rate as well as recognition rate is calculated from the expected characters and recognized characters for this subset using that specific window size. The recognition rate is tabulated in the corresponding cell in figure 4.9. The same process is repeated for each one of the 255 subsets and 4 window sizes. Therefore, for each document image being tested we can have 255 binary images as well as error rate/recognition rate when they are passed through ABBYY for a specific window size and we tabulate the corresponding recognition rate in the table shown in figure 4.9. Once all the cells are filled up, we search the highest recognition rate in the grid. Then the corresponding feature combination and window size of the highest recognition rate is determined. The same process is

Features Subset	Recognition Rate (%)			
	3x3	5x5	7x7	9x9
{1}	-	-	-	-
.....	-	-	-	-
{1,2}	-	-	-	-
.....	-	-	-	-
{1,2,3}	-	-	-	-
.....	-	-	-	-
{1,2,3,4,5,6}	-	-	-	-
.....	-	-	-	-
{1,2,3,4,5,6,7}	-	-	-	-
.....	-	-	-	-
{1,2,3,4,5,6,7,8}	-	-	-	-

Figure 4.9: Grid-based Method for Tabulating Recognition Rate- Subset vs. Window Size

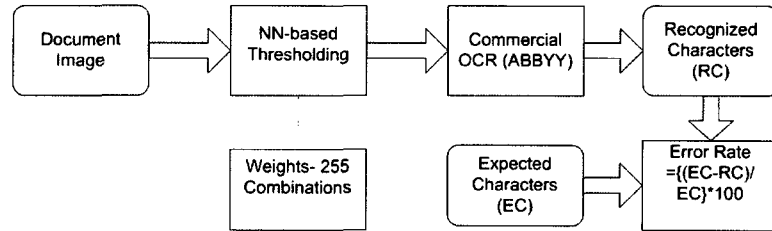


Figure 4.10: Test Setup

repeated for more testing images and obtain the best feature combination and window size for each image and then select the best among best combinations. This output is termed as optimal features subset. The flow chart of the test setup for optimal parameters selection is shown in figure 4.10.

4.3.2 Phase II-Efficient ANN-based Thresholding Method

In section 4.3.1 optimal feature selection method is described. The optimal features and it's corresponding window size in terms of the highest recognition rate is determined. The number of features in the optimal subset is used to define the efficient ANN structure. The efficiency of the proposed structure is defined in terms of the recognition rate, computational time and PSNR that will be

Table 4.1: Proposed Efficient ANN Structure

Input Units	n	3
Hidden Units	$(n + 1) * \frac{2}{3}$	2 (verified by DTREG [68])
Output Unit	1	1
Weights	$n * ((n + 1) * \frac{2}{3}) + (n + 1) * \frac{2}{3}$	8

discussed in section 4.4 and 4.4.4. The efficient ANN structure using the optimal features subset is shown in table 4.1. For example, the number of optimal features is n that determines the number of units in input layer. In this application the number of units in output layer is 1 (background=0 and object=1). Therefore, according to the DTREG [68] (a predictive modeling commercial software) [68], the number of hidden units in hidden layer will be $(n + 1) * \frac{2}{3}$. The weights are calculated from the input and output units using the relationship shown in table 4.1. Our simulation results show that the value of n is 3.

4.4 Simulation Results

A handful of binarization performance criteria [69] have been proposed, based on OCR performance criteria which is one of the most acceptable method. The binary image obtained from the ANN-based method is then passed to the commercial OCR for example ABBYY (version 7.0). The ABBYY system has its own image pre-processing tool which converts gray scale image into binary image and feeds them on for OCR [70]. In other words, original gray scale images without any binarization can be directly fed to ABBYY for recognition, as well as pre-binarized images. The results obtained from this method are regarded as a standard reference for comparison with other methods.

To test the performance of the proposed efficient ANN-based thresholding techniques a comparative study is carried out. In our experimentation, Otsu's global thresholding [37], Niblack [54], HMM [32] and Kittler[38] were utilized as benchmark for this comparison. To validate the optimal parameters, binary image obtained using the proposed technique will be compared to that obtained using Niblack [54] and Otsu [37] techniques in terms of yielding the character recognition rate.

4.4.1 Image Database

An image database is used for training and testing the proposed efficient ANN-based thresholding method and also to evaluate the efficiency of the proposed method. The database of total 120 images

was synthetically produced under different resolution and rotation. Total about 15000 characters belong to the image database. These images are categorized into non-uniform illumination, low contrast and composite image with complex background with different level of complexities. One image from each category shown in figure 4.2, 4.3 and 4.4 will be used to show the simulation results.

4.4.2 Training and Testing Results

The ANN-based local thresholding technique [7, 67] is applied to several images for preparing the training dataset as well as for testing. In that case 8 features are extracted for each pixel using different window size and save those features in different file for further use. About 20% images were used for generating training dataset using all 8 features and rest of them was used for testing purpose. Since 255 different combinations are available with 8 features, therefore for each testing image total $255 \times 4 = 1020$ different binary images are available that is obvious from the grid shown in figure 4.9. These images are then passed through the commercial OCR and the obtained recognized characters are compared to the expected characters to calculate the corresponding recognition rates. These recognition rates are tabulated in the grid shown in figure 4.9. The highest recognition rate in the grid represents the best feature combination.

Initially we have chosen one low contrast image of 654 characters shown in figure 4.2 for testing the proposed method. 3 subsets are found best among 255 subsets in terms of their performance and they have a very similar recognition rate (1 or 2 characters difference) in various window sizes. The original image without binarization is then passed through the commercial OCR ABBYY and the recognition rate is compared with the results obtained from the binary images using best 3 subsets. The same process is repeated for 2 more images and results are tabulated in table 4.2. Although we have obtained the recognition rate for window sizes 3x3, 5x5, 7x7 and 9x9, for comparison purpose we have shown the sample output in table 4.2 only for window size 5x5 rest of them is shown in appendix, because our work is based on the paper [67] that describes the ANN-based thresholding techniques and found 98% correct recognition rate at window size 5x5 using 8 features. Therefore, our intention is to show how the proposed method performs for the same window size. It is observed from table 4.2 that feature subsets {1, 2, 5}, {1, 2, 6} and {1, 5, 6} provides the best results in terms of recognition rate. The recognition rates of these 3 subsets in different window sizes are close to each other. Then from these three combinations, two combinations {1, 2, 5} and {1, 2, 6} are picked based on their average recognition rate. The same procedure is repeated for more images. The results are shown in table 4.3 for window size 5x5. Considering all types of images and their corresponding recognition rate the combination of three selected features "**pixel value (1),**

Table 4.2: Comparative Statement of Recognition Rates- Using Binary and Original Image

Image	Recognition Rate(%)- Binary Image			Recognition Rate (%)
Features Subset	{1, 2, 5}	{1, 2, 6}	{1, 5, 6}	ABBY (7.0)- Original Image
Health (Fig. 4.3)	99.79	99.79	99.58	99.37
Rail Road (Fig. 4.2)	99.85	99.54	99.54	99.24
Arnold (Fig. 4.4)	99.51	99.01	96.30	96.79
Average	99.72	99.45	98.07	98.47

Table 4.3: Comparative Statement of Recognition Rates-Features Combination

Image	Features-Recognition Rate(%)	
Features Subset	{1, 2, 5}	{1, 2, 6}
Niagra Falls (Fig. 4.6(a))	100	99.41
Chretien (Fig. 4.6(b))	99.77	99.53
George (Fig. 4.6(c))	99.52	99.28
Volcanoes (Fig. 4.7(c))	100	99.71
Cats (Fig. 4.7(d))	99.66	99.66
Average	99.79	99.52

mean (2) and entropy (5)” was found to be the best combination yielding the recognition rate of 99.79%.

The same process is applied to more testing images of various complex backgrounds for different window sizes using the best feature subset. For each window size the binary image is generated and is saved in a bit-map (BMP) file. The thresholding time is also recorded and is shown in table 4.4 for sample images and the corresponding binary image is passed through the ABBY to obtain the recognition rate. The character recognition rate is shown in table 4.5 for each window size. The computational time for processing one pixel centered on the different window size is shown in table 4.6. The processing time includes feature extraction and classification of the pixel into object or background. The processing time for each pixel and character recognition rate with respect to different window size is shown in figure 4.11 and 4.12 respectively.

From tables 4.5, 4.6 and figures 4.11 and 4.12, it is observed that window size does not play a significant role especially in character recognition rate. However, increase in window size results in higher computational costs hence slower the segmentation process. Therefore, comparing the computation costs and recognition rate, window size 3x3 is found to be the optimal in this application

Table 4.4: Window Size Vs. Thresholding Time (Sec)

Image	Image Size	Window Size- Time(Sec)			
		3x3	5x5	7x7	9x9
Arnold (Fig. 4.4)	689x723	6.66	7.66	9.11	11.15
Health (Fig. 4.3)	543x759	5.60	6.57	7.77	9.44
Rail (Fig. 4.2)	529x1191	8.38	9.91	11.62	14.10
Cat (Fig. 4.7 (d))	463x877	5.58	6.35	7.64	9.14
George (Fig. 4.6 (c))	693x712	7.01	8.00	9.47	11.37
Volcano (Fig. 4.7 (c))	492x886	5.96	6.85	8.25	9.84
Niagra Falls (Fig. 4.6 (a))	723x845	7.32	9.92	11.45	15.13
Chretien (Fig. 4.6 (b))	499x726	5.59	6.32	7.46	9.36
Health Canada (Fig. 4.7 (a))	479x660	5.13	5.51	6.29	7.52
Mercury (Fig. 4.7 (b))	538x799	6.43	7.94	8.19	10.33
Urban (Fig. 4.6 (d))	680x966	10.11	11.08	12.42	15.73
Insect (Fig. 4.5)	869x1084	14.40	16.02	18.21	22.43

using ANN-based local thresholding approach. This research indicates that 3 features used (Pixel value, Mean and Entropy) outperforms the 8 features utilized by Alginahi [7, 67] and 5 features utilized by the same author [63]. The processing time for 3, 5 and 8 features is tabulated in table 4.7 for an image of size 689x723 at window size 3x3.

Following the procedure described above, binary document image using proposed method of figure 4.2, 4.3 and 4.4 are shown in figure 4.13, 4.14 and 4.15 respectively.

4.4.3 Optimal Features Validation

To validate the optimal features and the window size a known binary document image (black text in white background) called base document is used. That document is added to the simple watermark image to the complex watermark image called watermarked image. In each case the watermarked image is thresholded using 3, 5 and 8 features. The difference between the thresholded image and base document is the noise. Peak-Signal-to-Noise ratio (PSNR) is used to measure the noise in each case. PSNR is defined as follows:

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \|I(i, j) - K(i, j)\|^2 \quad (4.2)$$

Table 4.5: Window Size Vs. Correct Recognition Rate (%)

Image	Total Chars	Window Size- Misclassification			
		3x3	5x5	7x7	9x9
Arnold (Fig. 4.4)	405	0	3	1	3
Health (Fig. 4.3)	476	2	2	2	0
Rail (Fig. 4.2)	654	0	1	1	2
Cat (Fig. 4.7 (d))	585	0	1	1	1
George (Fig. 4.6 (c))	414	0	1	1	1
Volcano (Fig. 4.7 (c))	384	0	1	0	0
Niagra Falls (Fig. 4.6 (a))	339	1	0	0	0
Chretien (Fig. 4.6 (b))	426	3	0	1	0
Health Canada (Fig. 4.7 (a))	383	0	0	0	0
Mercury (Fig. 4.7 (b))	301	0	0	0	0
Urban (Fig. 4.7 (d))	692	0	0	0	0
Insect (Fig. 4.5)	487	4	20	6	8
Total	5546	10	29	13	15
Correct (%)	100	99.82	99.48	99.77	99.73

$$PSNR = 20 * \log \frac{MAX_I}{\sqrt{MSE}} \quad (4.3)$$

where $I(i, j)$ is the base document and $K(i, j)$ is the thresholded image of size $m \times n$ using proposed method, and MAX_I is the maximum possible value of an image. For gray scale image it is 255. PSNR is tabulated in table 4.8 for 4 sample images. Sample base document, watermark image, watermarked document image and thresholded image using the proposed optimal features are shown in figures 4.16-4.29.

Table 4.6: Window Size Vs. Process Time for each pixel

Window Size	Process Time (μs)
3x3	13.5
5x5	15.7
7x7	18.5
9x9	22.6

Table 4.7: Number of Features Vs. Process Time (Sec)

No. of Features	Process Time (Sec)
3 (Proposed)	6.60
5 (Alginahi [63])	8.47
8 (Alginahi [7, 67])	9.03

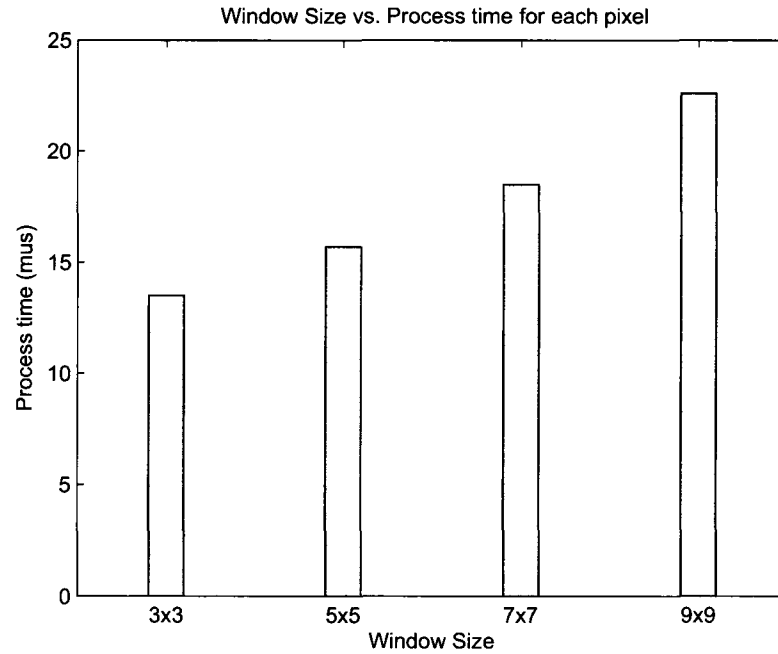


Figure 4.11: Window Size Vs. Process Time (μs) for each pixel

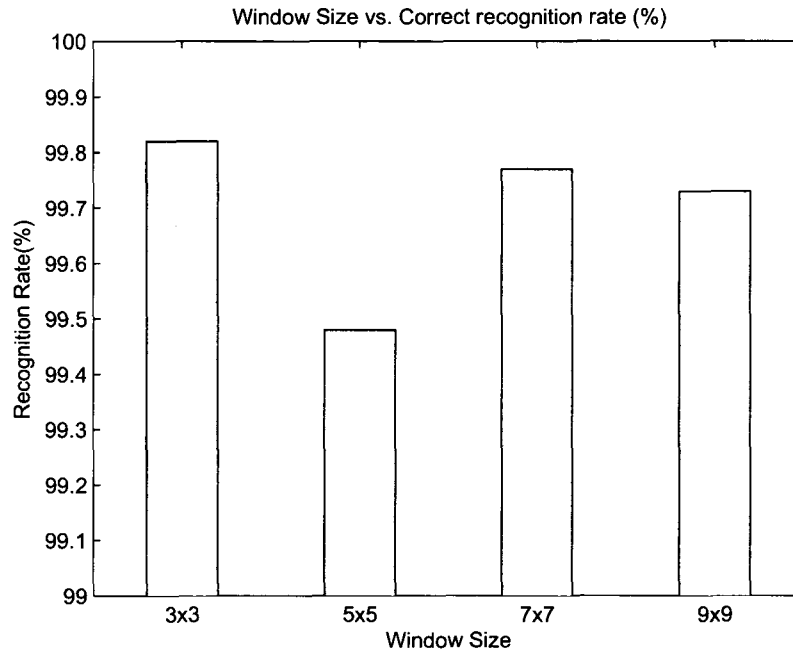


Figure 4.12: Window Size Vs. Character recognition rate (%)

The Underground Railroad

In the days before and during the American Civil War, Ontario served as the final stop on the underground railroad, a network of secret routes and safe houses that allowed enslaved African Americans to escape to freedom in Canada.

Walk in the footsteps of history along the African Canadian Heritage Route from Windsor, where you can visit John Freeman Wall's 1846 log cabin, that served as a terminal on the Underground Railroad. For a further window into the past, walk among the artifacts and images at the Amherstburg's North American Black Historical Museum. Stroll the streets of North Buxton, Canada's first Black settlement - home to many historic buildings and a museum that recounts the area's proud story of early growth and self-sufficiency.

Figure 4.13: Binary Image Using Proposed Method of figure 4.2

Health is defined as a state of complete physical, social and mental well-being, and not merely the absence of disease or infirmity. Within the context of health promotion, health has been considered less as an abstract state and more as a means to an end which can be expressed in functional terms as a resource which permits people to lead an individually, socially and economically productive life. Health is a resource for everyday life, not the object of living. It is a positive concept emphasizing social and personal resources as well as physical capabilities.

Figure 4.14: Binary Image Using Proposed Method of figure 4.3



In early 2003, Californians pointed fingers as their state struggled with a \$38 billion budget deficit and a continuing energy crisis. Republicans set their sights on

Democratic Gov. Gray Davis, attempting to make him the second governor in U.S. history to be recalled. On October 7, the majority of voters decided to oust Davis, then chose a successor from among 135 candidates. One of Hollywood's own took Davis place bodybuilder-turned-actor Republican Arnold Schwarzenegger.

Figure 4.15: Binary Image Using Proposed Method of figure 4.4

Table 4.8: PSNR Measurement for Different Parameters

Image	Image Size	No. of Features	No. of Error Pixels	PSNR-dB
Fig. 4.26	426x848	3 (Proposed)	4	49.5575
		5 [63]	48088	8.7577
		8 [7, 67]	22077	12.1387
Fig. 4.27	426x848	3 (Proposed)	0	∞
		5 [63]	37187	9.8741
		8 [7, 67]	18230	12.9702
Fig. 4.28	426x848	3 (Proposed)	3992	19.5661
		5 [63]	22079	12.1383
		8 [7, 67]	18649	12.8715
Fig. 4.29	659x638	3 (Proposed)	1573	24.2698
		5 [63]	34527	10.8555
		8 [7, 67]	29443	11.5472

Digital watermarking is the process of embedding information into a digital signal. The signal may be audio, pictures or video, for example. If the signal is copied, then the information is also carried in the copy.

Figure 4.16: Sample Base Document 1

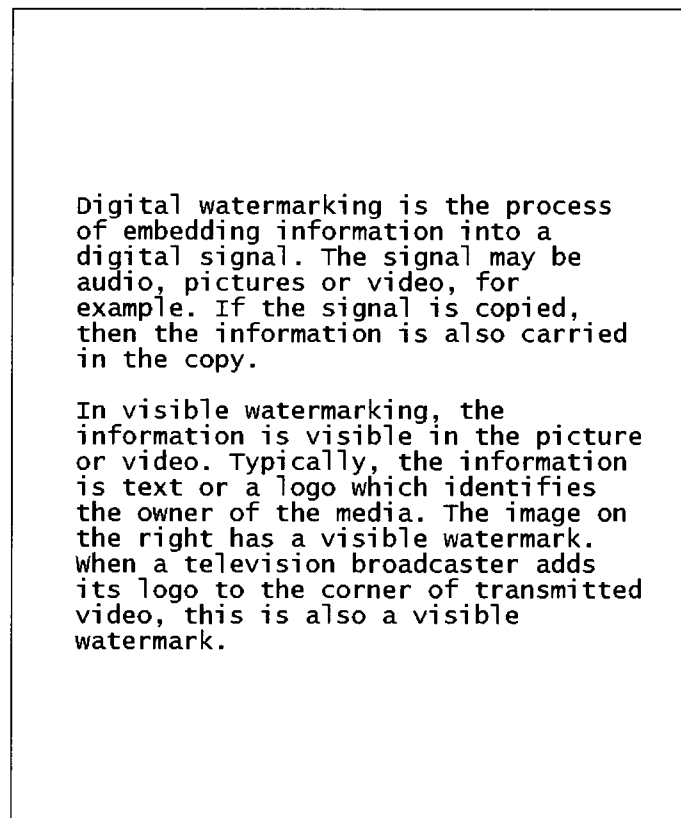


Figure 4.17: Sample Base Document 2

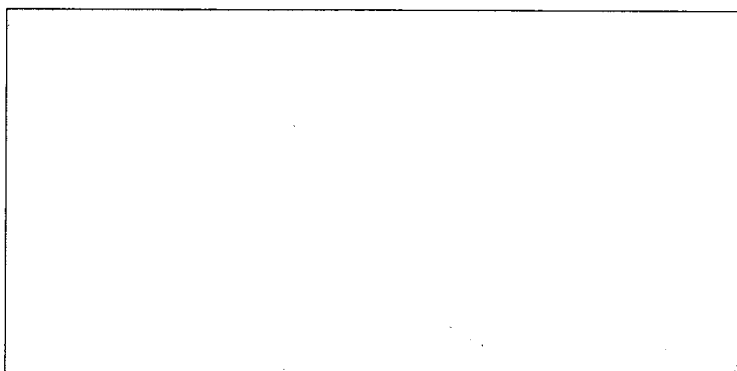


Figure 4.18: Sample Watermark 1

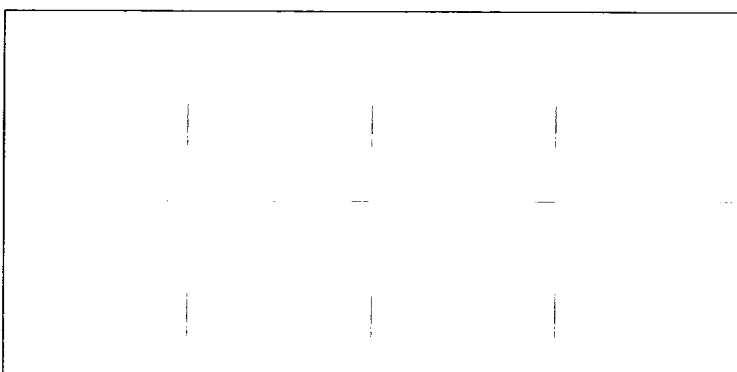


Figure 4.19: Sample Watermark 2

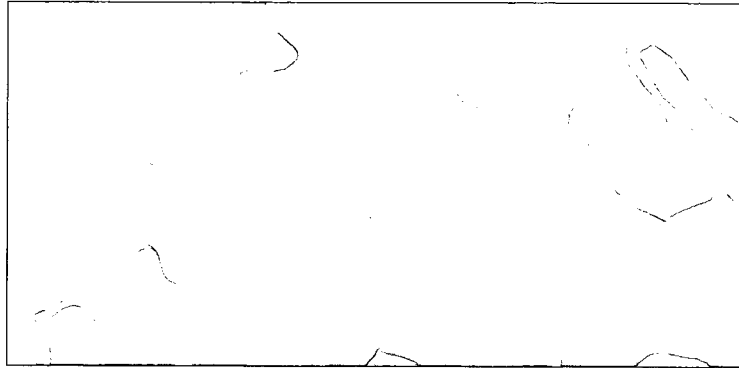


Figure 4.20: Sample Watermark 3

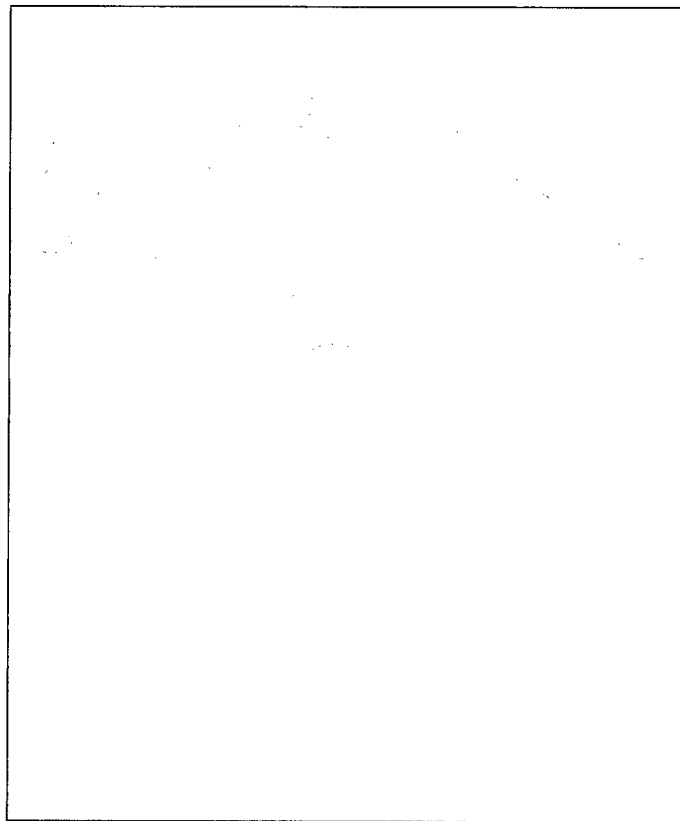
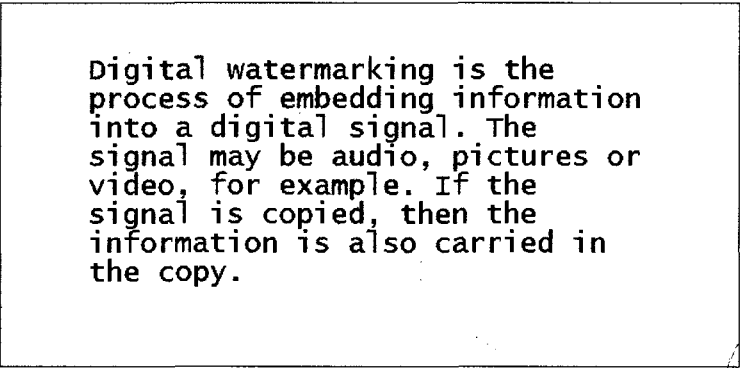
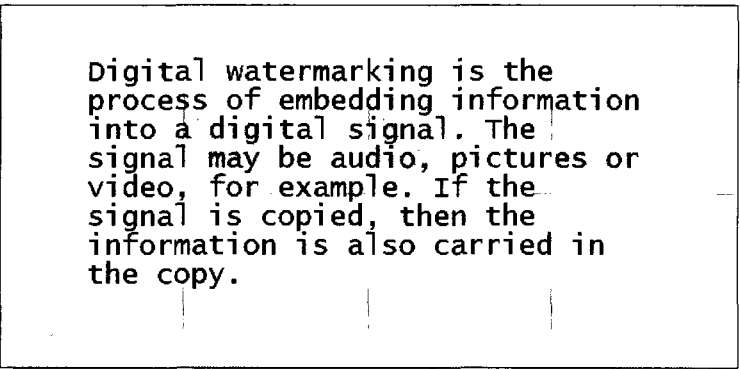


Figure 4.21: Sample Watermark 4



Digital watermarking is the process of embedding information into a digital signal. The signal may be audio, pictures or video, for example. If the signal is copied, then the information is also carried in the copy.

Figure 4.22: Sample Document Watermarked Image 1



Digital watermarking is the process of embedding information into a digital signal. The signal may be audio, pictures or video, for example. If the signal is copied, then the information is also carried in the copy.

Figure 4.23: Sample Document Watermarked Image 2

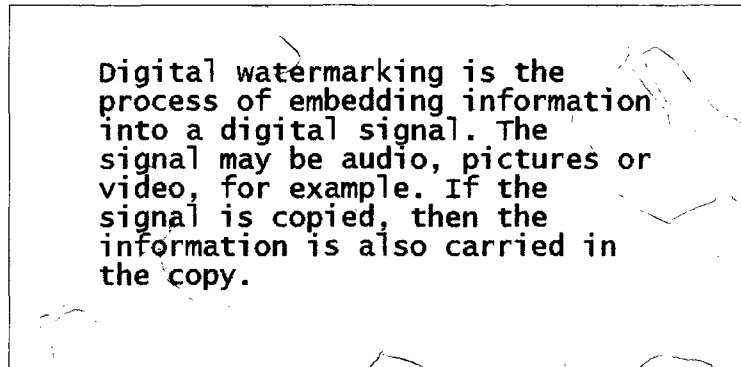


Figure 4.24: Sample Document Watermarked Image 3

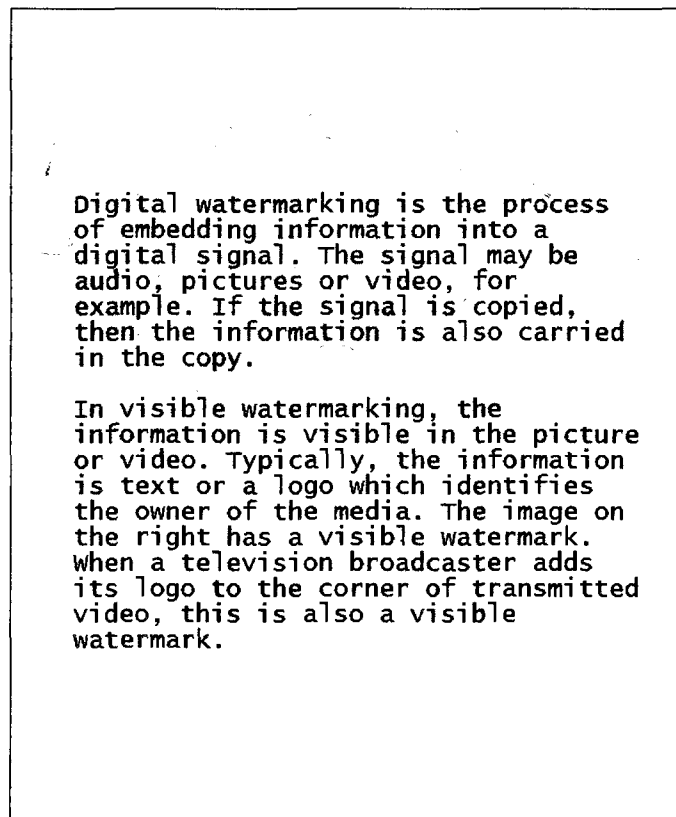


Figure 4.25: Sample Document Watermarked Image 4

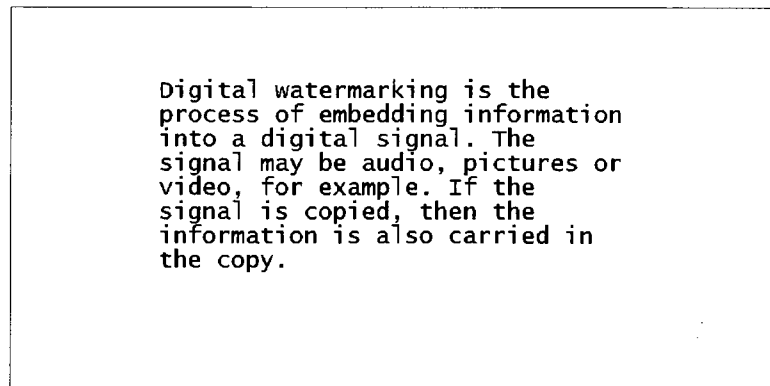


Figure 4.26: Binary Image 1 Using Proposed 3 Features

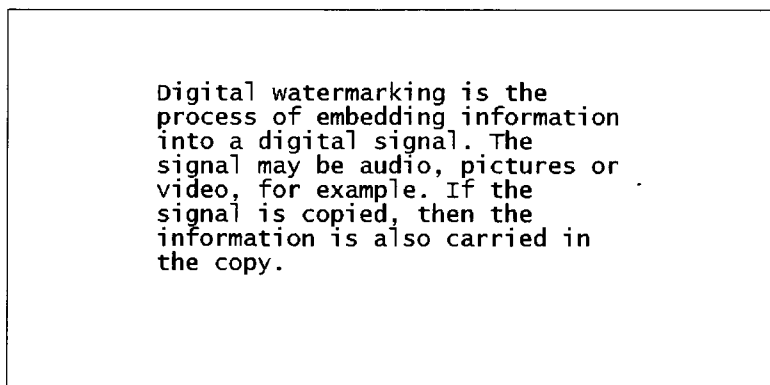


Figure 4.27: Binary Image 2 Using Proposed 3 Features

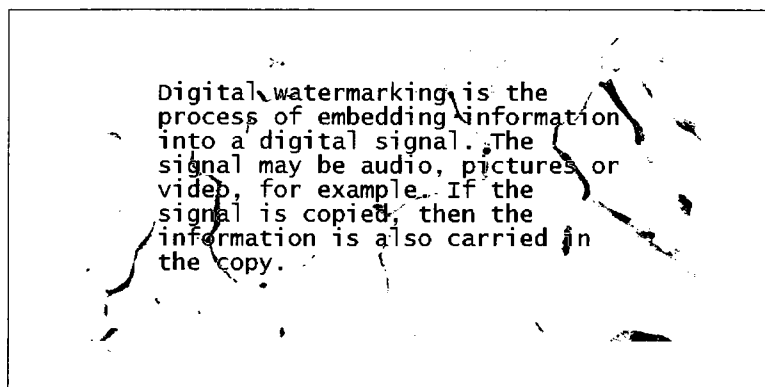


Figure 4.28: Binary Image 3 Using Proposed 3 Features

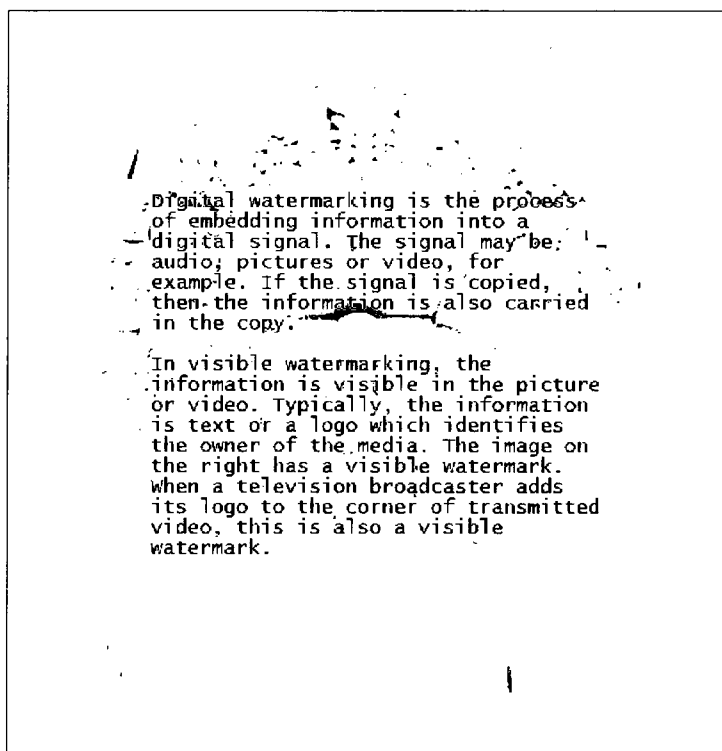


Figure 4.29: Binary Image 4 Using Proposed 3 Features

Table 4.9: Recognition Rate Comparison- Proposed, Niblack [54] and Otsu [37]- (Figure 4.4)

Techniques	Misclassification (chars)	Correct Recognition Rate (%)
Proposed	0	100
Niblack [54]	11	97
Otsu [37]	59	85

4.4.4 Performance Evaluation of the Proposed Method

To evaluate the performance of the proposed efficient ANN-based thresholding method, the obtained binary image using the optimal features at optimal window size is passed through ABBYY and recognized characters output is generated. Sample OCR output from the binary image and the expected output is shown in figure 4.30 for an image shown in figure 4.4. To validate the performance of the proposed combination the output from the original image is compared to the expected output that is shown in figure 4.31. From these two figures 4.30 and 4.31, it is obvious that the proposed method shows better performance in terms of recognition rate. OCR output of two other binary images 4.13 and ?? using proposed method are shown in figure 4.32 and 4.33 respectively. These outputs are compared to expected output and indicated the error occurred.

The output of the proposed ANN-based thresholding technique using 3 optimal features at window size 3x3 is also compared to Otsu [37] and Niblack [54] method. The binary image and their corresponding OCR outputs of Otsu [37] and Niblack [54] are shown in figures 4.34, 4.35 and table 4.9 for a document image of 405 characters (figure 4.4).

Another binary image obtained using proposed features is compared with Huang [32] based on HMM, Otsu [37] and Kittler [38] method that is shown in figure 4.36. From this figure it is obvious that Kittler [38] shows better performance than HMM [32]. However Kittler [38] has some broken characters in the binary image which is not present in the binary image obtained using proposed method. Therefore, the proposed method shows better performance than any one methods described in the literature. Then the proposed method is evaluated based on three different categories of document image with low contrast, non-uniform illumination and composite with complex background in terms of recognition rate for unknown base documents and PSNR for known base documents and the result is shown in figure 4.10. Finally, the overall simulation results for around 15000 characters is shown in table 4.11.

<ul style="list-style-type: none"> in early 2003, Californians pointed fingers as their state struggled with a \$38 billion budget deficit and a continuing energy crisis. Republicans set their sights on Democratic Gov. Gray Davis, attempting to make him the second governor in U.S. history to be recalled. On October 7, the majority of voters decided to oust Davis, then chose a successor from among 135 candidates. One of Hollywoods own took Davis place bodybuilder-turned-actor Republican Arnold Schwarzenegger. 	<ul style="list-style-type: none"> In early 2003, Californians pointed fingers as their state struggled with a \$38 billion budget deficit and a continuing energy crisis. Republicans set their sights on Democratic Gov. Gray Davis, attempting to make him the second governor in U.S. history to be recalled. On October 7, the majority of voters decided to oust Davis, then chose a successor from among 135 candidates. One of Hollywoods own took Davis place bodybuilder-turned-actor Republican Arnold Schwarzenegger
Recognized	Expected

Figure 4.30: OCR output: From Binary Image 4.15 (Proposed) vs. Expected

Table 4.10: Performance Evaluation of the Proposed Method in Different Categories

Category	Unknown Base		Known Base
	Chars (6996)	Recognition Rate (%)	PSNR (dB)
Low Contrast	1463	99.79	49.56
Non-uniform Illumination	2192	99.86	19.57
Composite with Complex Background	3341	99.76	24.27

4.5 Conclusions

An exhaustive search was conducted to develop an efficient thresholding technique based on the work reported in [7, 63, 67]. In this thesis, a method is devised to find the optimal parameters and optimal window size in ANN-based thresholding technique to segment the non-uniform and composite document images. Three features **pixel value**, **mean** and **entropy** and window size 3x3 are found to be optimal in this technique in terms of PSNR, computational costs and correct recognition rate. To measure the performance of the proposed method, the segmented image obtained by the optimal features was compared with HMM [32], Otsu [37], Niblack [54] and Kittler [38]- based segmented image. For known base documents and watermark images, the noise is measured in terms of PSNR from the segmented image obtained using proposed 3 optimal features, 5 features [63] and 8 features [67] and is compared. The result shows that the proposed method outperforms the reference methods. Then the optimal features were used to segment the same set of document image files used in

<ul style="list-style-type: none"> in early 2003, Californians pointed fingers as their state struggled with a \$38 billion budget deficit and a continuing energy crisis. Republicans set their sights on Democratic Gov. Gray Davis, attempting to make him the second governor in U.S. history to be recalled. On October 7, the majority of voters decided to oust Davis, then chose a successor from among 135 candidates. One of Hollywood's own took Davis' place: bodybuilder-turned-actor Republican Arnold Schwarzenegger. 	<ul style="list-style-type: none"> In early 2003, Californians pointed fingers as their state struggled with a \$38 billion budget deficit and a continuing energy crisis. Republicans set their sights on Democratic Gov. Gray Davis, attempting to make him the second governor in U.S. history to be recalled. On October 7, the majority of voters decided to oust Davis, then chose a successor from among 135 candidates. One of Hollywood's own took Davis' place: bodybuilder-turned-actor Republican Arnold Schwarzenegger.
ABBY	Expected

Figure 4.31: OCR output: From Original Image vs. Expected

<p>The Underground Railroad</p> <p>in the days before and during the American civil War, Ontario served as the final stop on the underground railroad, a network of secret routes and safe houses that allowed enslaved African-Americans to escape to freedom in Canada.</p> <p>Walk in the footsteps of history along the African Canadian Heritage Route from Windsor, where you can visit John Freeman Walls' 1846 log cabin, that served as a terminal on the Underground Railroad. For a further window into the past, walk among the artifacts and images at the Amherstburg's North American Black Historical Museum. Stroll the streets of North Buxton, Canada's first Black settlement - home to many historic buildings and a museum that recounts the area's proud story of early growth and self-sufficiency.</p>	<p>The Underground Railroad</p> <p>In the days before and during the American civil War, Ontario served as the final stop on the underground railroad, a network of secret routes and safe houses that allowed enslaved African-Americans to escape to freedom in Canada.</p> <p>Walk in the footsteps of history along the African Canadian Heritage Route from Windsor, where you can visit John Freeman walls' 1846 log cabin, that served as a terminal on the Underground Railroad. For a further window into the past, walk among the artifacts and images at the Amherstburg's North American Black Historical Museum. Stroll the streets of North Buxton, Canada's first Black settlement - home to many historic buildings and a museum that recounts the area's proud story of early growth and self-sufficiency.</p>
Recognized	Expected

Figure 4.32: OCR output: From Binary Image 4.13 (Proposed) vs. Expected

[7] and applied to a commercial OCR ABBYY and 99.25% recognition rate is achieved for around 15000 characters. Our quantitative study shows that recognition rate obtained using the proposed method is much higher than the methods found in the literature.

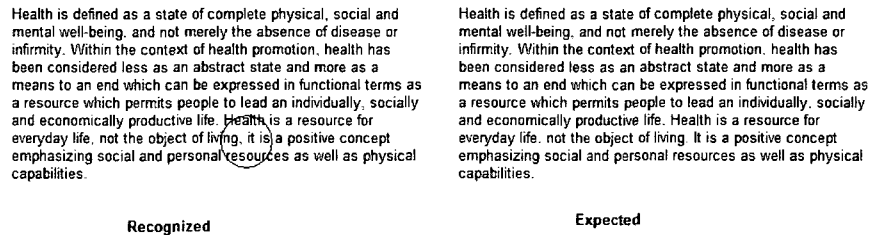


Figure 4.33: OCR output: From Binary Image 4.14 (Proposed) vs. Expected

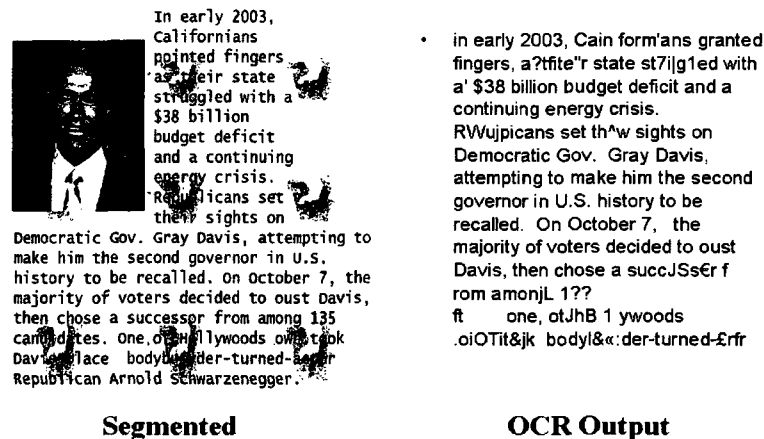


Figure 4.34: OCR output: From Binary Image (Using Otsu [37])

Table 4.11: Performance of the Proposed Method Using Optimal Features

Technique	Tested Chars	Recognition Rate (%)
Proposed, 3 features	15000	99.25
ABBY [7]- Original Image	14600	96
Alginahi [7], 8 features	14600	98
Alginahi [63], 5 features	Not Mentioned	98.3

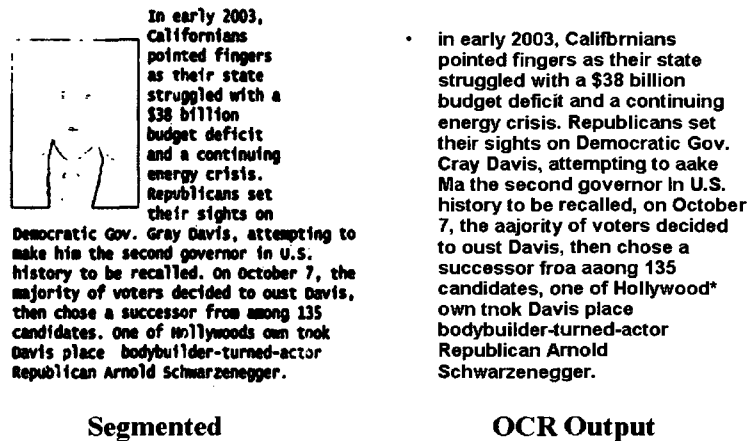


Figure 4.35: OCR output: From Binary Image (Using Niblack [54])



Figure 4.36: Segmentation Performance Comparison: Proposed, HMM[32], Kittler [38] and Otsu [37]

Chapter 5

Edge-based Thresholding and Quality Process Control

5.1 Introduction

Two part (body and cap) gelatin capsules are sensitive to several common flaws caused by the equipment problems and inconsistencies of the manufacturing process that make them unacceptable for shipment and marketing. These flaws are incorrect size or color, dents, cracks, holes, bubbles, strings, dirt, double caps, missing caps and so on. Some of these are called critical defects. A hole or crack is one of the critical defects and once the capsule is filled it is possible to spill out the product and can cause serious problem if it reaches to the end-user. The area of bubble is much thinner and can easily turn into a hole. Dents have irregular shape and cause failure at the filling stage. Incorrect color, size and double caps can cause serious problems if they reach to the end-user, resulting in possible product recall. Dirt and strings are cosmetic flaws and do not affect the capsule, however it implies poor quality control of production. Typical defects are shown in figure 5.1. In this thesis we focus on solving a pharmaceutical manufacturing problem.

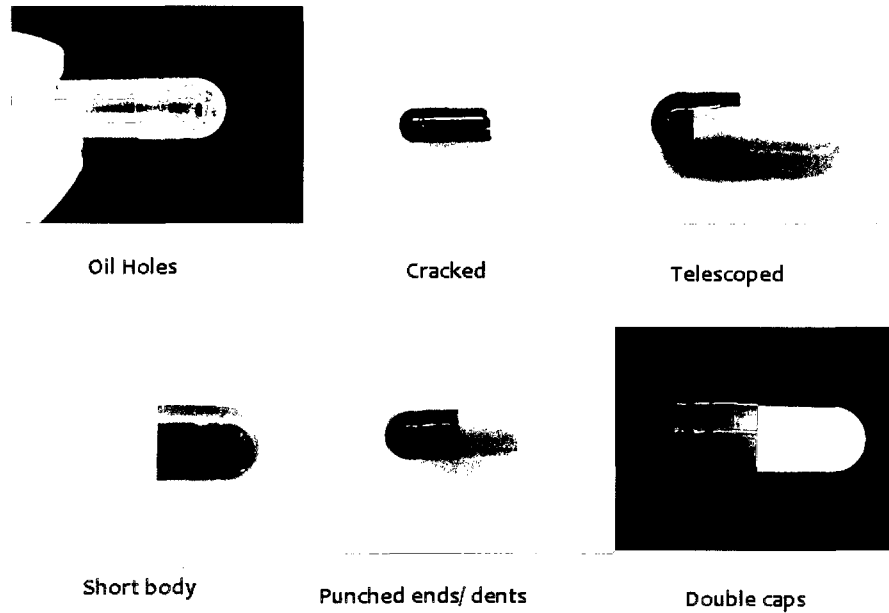


Figure 5.1: Typical defects (Source: Pharmaphil Inc., Windsor, ON)

5.2 Existing Quality Control and Criticisms

Pharmaphil Inc. a Windsor, ON based pharmaceutical industry manufactures two-part gelatin capsules. Currently human inspectors evaluate the quality of the capsules manufactured at Pharmaphil Inc. This process involves a worker monitoring a conveyor belt of capsules and removing any defective product that is shown in figures 5.2 and 5.3. The capsules are then collected in bins, each of which contains 1,00,000 capsules. Manual inspection has some advantages, such as a human inspector can provide immediate solution for quality control as they are trained to look for a set of various flaws in product line. The adaptability of human inspector to changing environments and different products is incomparable to any automated system. However, the drawback of manual inspection costs more. First of all, manual inspection is not consistent and it is unreasonable to expect that all human inspectors will provide the same inspection output. Secondly, human workers can not keep their attention at same degree at all times for the duration of a given shift yielding inaccurate inspection results. Given that situation, when the inspectors find a defective capsule in one bin, the company discards the entire bin of capsules. This process results in a loss of revenue, a waste of manufacturing time, increase in manufacturing costs and uncertainty about the quality of the products the company sends to the market.

Therefore, developing low-cost, high throughput quality process controller is a prime need to

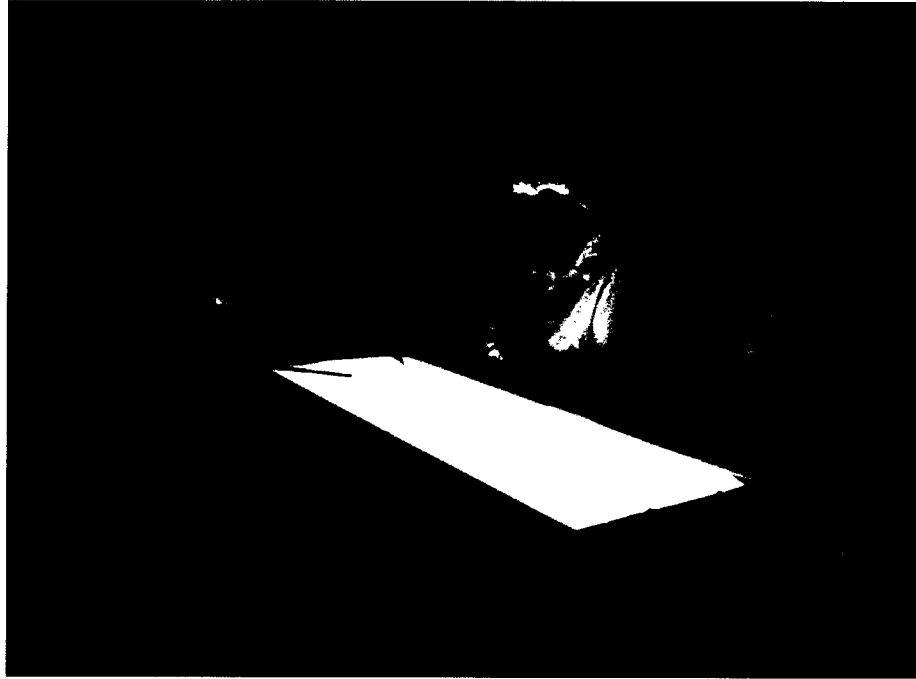


Figure 5.2: Current capsule sorting system

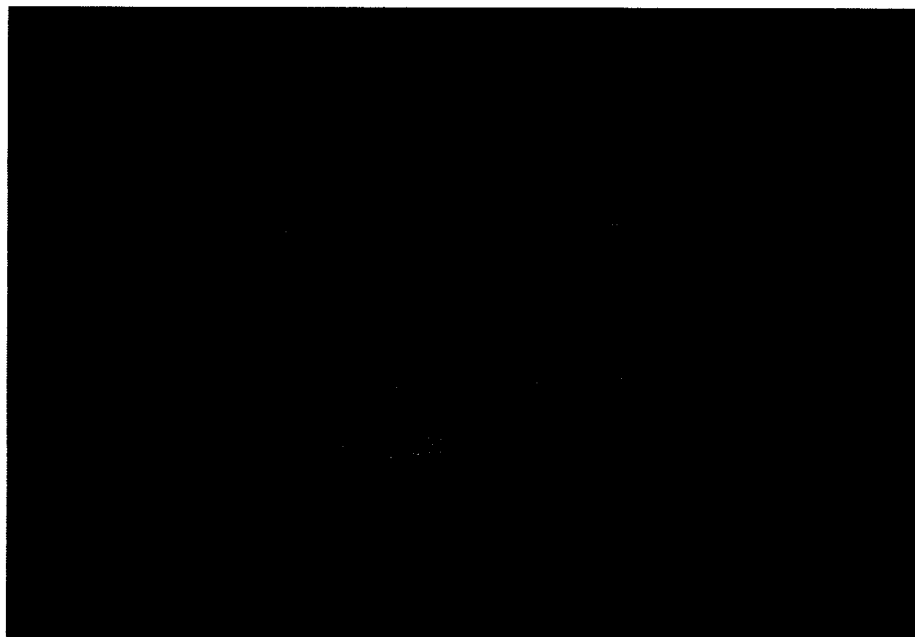


Figure 5.3: Current capsule sorting system- Conveyor belt

improve the productivity of the company, to improve the quality of their manufactured capsules and enable the company to better compete globally. To fulfill these objectives of the Pharmaphil Inc., a computationally efficient image processing technique was developed in addition to the low-cost hardware systems to perform the quality inspection of the capsules in collaboration with the Electrical and Computer Engineering Department at the University of Windsor. The main aim was to develop a computer vision system capable of isolating and inspecting capsules at the rate of 1000 capsules per minute with over 95% accuracy of detecting defects of size 0.2 mm and larger with the ability to detect holes, cracks, dents, bubbles, double caps and incorrect colour or size.

5.3 Current State-of-the-art Solutions

A computer vision based system controller requires one time establishment cost and possible maintenance costs. There are few capsule sorting machines with quality process control settings available in the market at a cost of over \$500,000 per machine [71]. DJA- PHARMA, a division of Daiichi Jitsugyo America has Capsules Visual Inspection System CVIS-SXX-E that inspects uni-color and bi-color hard gelatin capsules of size 1 through 5 and these with printed marks except both body and cap are transparent and dark color [72]. The inspection rate of this system is 1700 to 2500 capsules per minutes. Viswill has a Capsules Video Inspection Systems named CVIS-SXX manufactured in Japan that uses high-resolution CCD line sensor cameras and is capable of detecting a minimum size flaws of 100 μm while Eisai Machinery USA has similar systems named CES-50, 100, 150 with an inspection rate of 800, 1600 and 2500 respectively. A France based company called ProdiTec has pharmaceutical industrial vision system called InspeCaps-150 that inspects hard capsules of size 0 to 5 at the rate of 2000 capsules per minutes. Therefore, although automated systems are available, they are very expensive and very few of them dealing with transparent capsules which is the case in the application at Pharmaphil Inc.

The main objectives of this project is to develop a working prototype that proves the concept of the capsule inspection system. A computationally efficient image processing techniques is a required addition to the low-cost hardware systems to develop the quality inspection systems of the capsules.

5.4 Capsule Inspection System- OptiSorter

In this pharmaceutical application, the capsule inspection system requires that every capsule is fully inspected to meet accuracy of defects of size 0.2 mm and larger. In order to recognize the defects

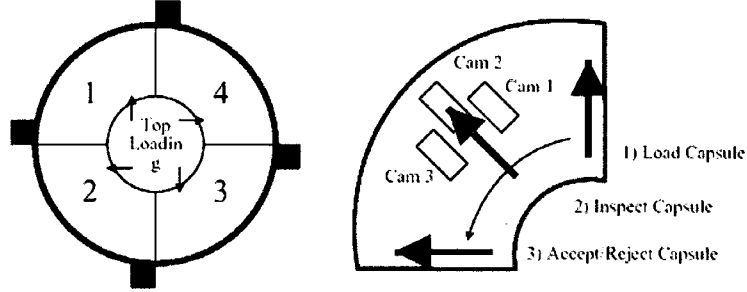


Figure 5.4: Image acquisition system

with a vision system, full 360° imaging should be acquired from multiple image sensors. The system must be able to verify the capsule color and size including other defects satisfying the throughput at a rate of 1000 fully inspected capsules per minute. The system must also be flexible enough to adjust for various sizes and colors of capsules, and accommodate future updates in image resolution and more thorough inspection algorithms [71].

The image acquisition and processing components are placed into an existing capsule sorter called OptiSorter which fixes the capsules individually in solid holders allowing full 360° visibility of the capsule. The system controller allows the vision system to interact with the mechanical system and coordinating the triggering of the cameras as well as controlling the accepting and rejecting mechanisms of the capsules.

To meet the system throughput, four identical inspection stations operate in parallel on a single mechanical system. A set of three USB cameras and a host PC belongs to each quadrant. Figure 5.4 shows the image acquisition system. It is great advantageous to use four stations that provides enough safeguards against system failure so that in any case if one station fails remaining stations can be operational properly [73].

Most of the vision system uses 3 cameras to cover full 360° views of a capsule, 120° separation of each camera from the other. However, it is recommended to use one more camera on top and the rationale will be discussed in conclusion of this chapter. Figure 5.5 shows the acquired images of capsules from four different angles.

5.5 Proposed Methodology

Where visual inspection is the main concern, automatic computer image analysis may provide a solution to these problems [74]. Therefore, it is required to know the different parts of the capsule.

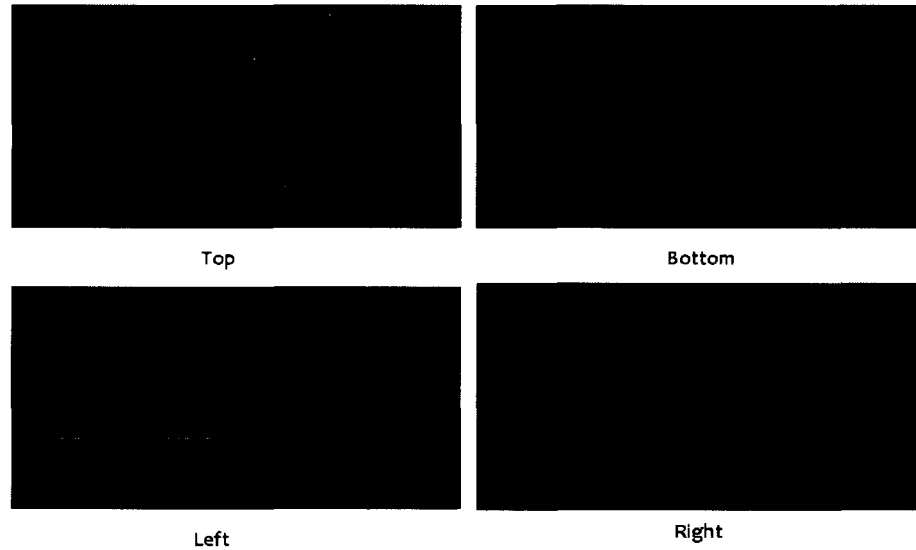


Figure 5.5: Capsule image- Top, Bottom, Right and Left camera view

A typical gelatin capsule image (size 0) is shown in figure 5.6. In these pharmaceutical applications, various capsules parameters are required to be evaluated and analyzed. For example, capsules dimension (length and width), body and cap lengths, body and cap diameter, intersection of body and cap (closed joined lengths) are the important parameters for quality inspection that determines the wrong size, missing cap, double cap and some of dents. The block diagram of the proposed quality inspection system is shown in figure 5.7.

5.5.1 Preprocessing

Background noise, unwanted reflections from poor illumination and geometric distortions are very common noise from the acquisition system during the capturing of the images of capsules by the digital cameras that makes the image segmentation very much sensitive. Image enhancement is a way of solving these problems as well as restores the image to improve overall clarity by correcting geometric distortions from the acquisition systems. Image enhancement techniques are used as preprocessing tools- a series of image-to-image transformations. It does not add more information but it may help to extract it. Visually it is hard to state the level of enhancement. Mostly, we are concerned with finding whether the enhancement has improved our ability to differentiate the object from background. This difficulty justifies the need for quantitative measures to quantify the performance of the enhancement techniques. Computational cost and global variance are proposed

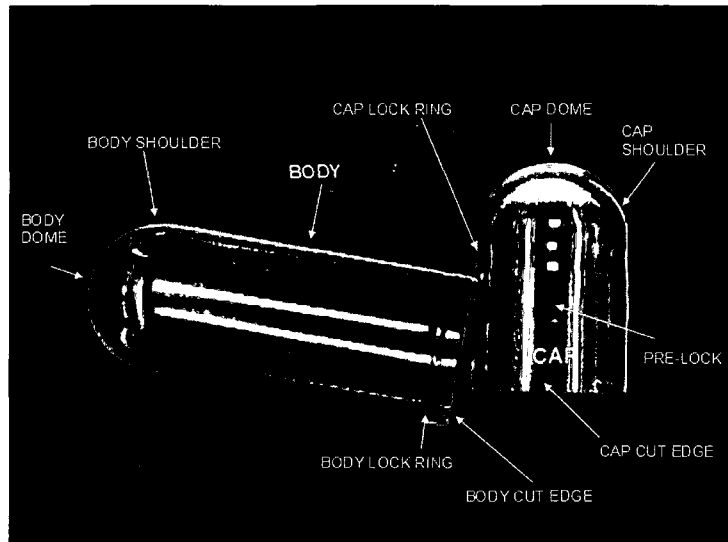


Figure 5.6: Typical capsule image (Source: Pharmaphill Inc.)

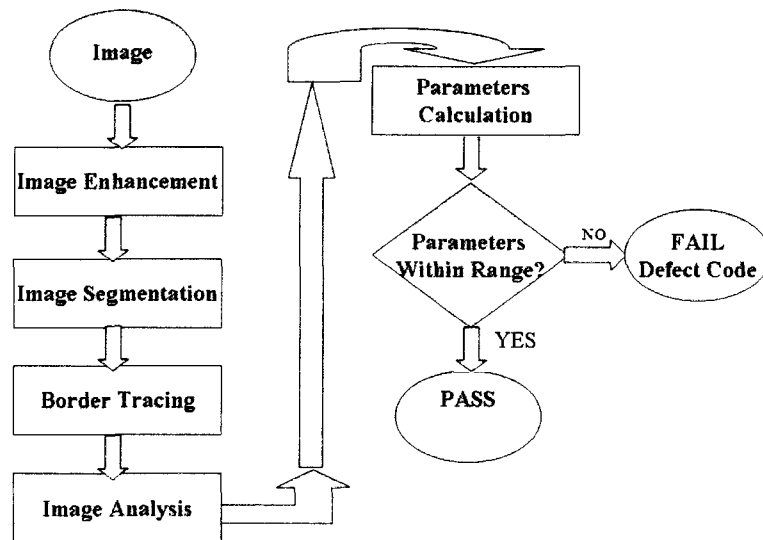


Figure 5.7: Block diagram of the proposed image processing system

as a quantitative measure in this application for each method and finally the best method for this application is selected both in terms of execution time and global variance. Global variance is defined as follows:

$$\sigma^2 = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} (g_{mn} - \bar{g})^2 \quad (5.1)$$

with an image size $M \times N$ and mean pixel value \bar{g} . It is defined as:

$$\bar{g} = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} g_{mn} \quad (5.2)$$

5.5.2 Image Segmentation

Image segmentation is an important first step for a variety of image analysis and visualization tasks [75]. Lots of image thresholding techniques are reviewed in the chapter 3. Some of the issues that make image segmentation difficult are-

- a) The image capturing process normally has inherent limitations which result in noisy, less than perfect images,
- b) Sometimes the sensitivity of the image data to solve a given task is very high, like the case with transparent capsule imaging, the high frequency information in them is often distorted resulting in fuzzy, non-reliable edges,
- c) From image to image even in the same modality, the shape of the same structure can vary,
- d) The gray scale values and their distributions vary from image to image even in a single application. Pixels corresponding to a single class may exhibit different intensities between images or even within the same image.

The main issues related to segmentation involve choosing good segmentation algorithms, measuring their performance, and understanding their impact on the capsule inspection system. But the segmentation does not perform well if the gray levels of different objects are quite similar which is the case in this application. Because capsules are clear, transparent that is overlapped with background gray levels. ANN-based thresholding technique is applied in this application to separate the capsule from the acquired image and results will be shown in results section.

Through observation one can notice that the values of pixels around the boundary changes abruptly even in case of complex background. In that case simple edge detector can easily distinguish the boundary of the objects from an inhomogeneous background [76, 77]. In edge-based segmentation techniques, selection of the edge detector is an important issue. In terms of time, quality and the requirement of the segmentation, Sobel edge detector is recommended in this application

because of its superior noise-suppression characteristics. Since Otsu [37] is the most referenced global thresholding method, it is used here to localize the edges after edge detection.

Dimension of the capsule as well as cap and body lengths are the important parameters for detecting the size, missing cap and double cap. Some dents can be detected from the size of the capsule. But most of the dents are small that do not even change the shape of the two ends of the capsule. These types of dents are very hard to detect that makes the job challenging. From the dented capsules, it was found almost 90% dents are in either cap end or body end. Although sometimes it does not make any difference in the length or width, but in most cases, the radius of these two ends differs much. This observation motivated us to come up with an idea to measure the radius of both ends and compare the measured radius with the expected radius. The proposed steps are given below-

- a) Divide the capsule image into 3 parts. The top and bottom parts are very similar to the dome (half circle) and they are approximated to a half circle using least square circle fitting. To do that, at first an edge-based segmentation is used to locate the transition between background and edges of the capsules and mathematical morphological operators are used to smooth the edges.
- b) Using suitable border-tracing algorithm [78] the border points of both top and bottom parts are obtained and these border points are approximated to a half-circle and the radius of each part and two end points of the half circle are identified. The first decision whether it is a defect free or defective capsule comes from the radius. If the radius is out of the ideal range, it is defective.
- c) From the top and bottom part half circles we have a total of 4 end points and 2 points in each side are connected together that form an approximated capsule image. The flowchart of the proposed steps to extract the capsule image from the captured image is shown in figure 5.8.

Once the approximated capsule image is extracted the next step is to analyze the capsule region to detect the defects. At this stage, each part of the capsule for example top (left), bottom (right), left lock and right lock are segmented using edge-based segmentation followed by morphological smoothing to calculate the different capsule parameters. These calculated parameters are compared with the standard parameters and if any of them goes out of range, it is concluded that the capsule is defective. Once the parameters are in the range, connected component labeling is applied and counts the number of pixels in each region to check the holes, bubbles or cracks. The number of pixels allowed to be a defect free capsule can vary according to the requirements of the industry and the proposed system has the flexibility to take this value.

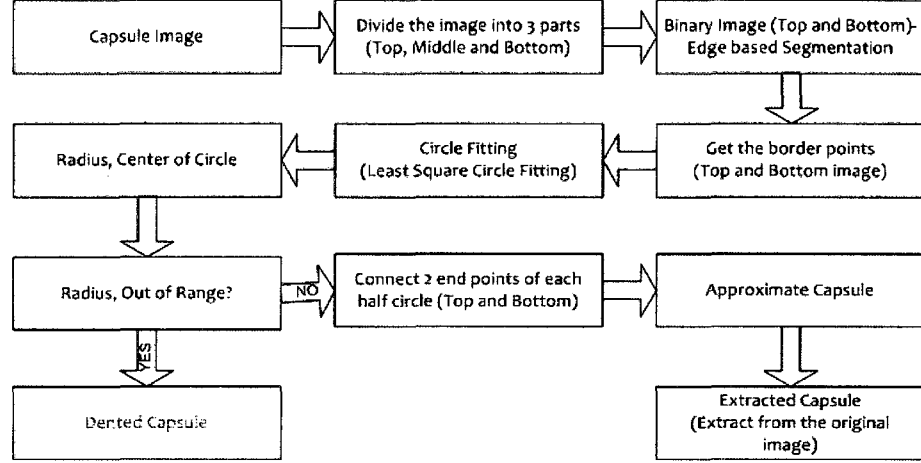


Figure 5.8: Flow chart of the proposed capsule extraction method

5.5.3 Border Tracing

The border tracing algorithm is used to extract the contour of the objects from an image. This section is based on a reference [78, 79] which gives an excellent overview of simple border tracing algorithm using an efficient table for calculating the neighborhood coordinates. The first step of border tracing is to search the image from top-left corner until a pixel of a new region is found. This pixel is labeled as P_0 with the coordinate (I_1, J_1) and called starting point. In this application the starting point is verified by checking it's neighboring points whether it is noise pixel or a pixel of the region of interest. The next steps to find the border points are as follows:

1. A variable *dir* is initialized to 7 to detect the border points in 8-connectivity that stores the direction of the previous move along the border from the previous border element to the current border element.
2. Search the 3x3 neighborhood of the current pixel in an anti-clockwise direction. The point immediately preceding (I_1, J_1) is called the first neighbor and labeled as (ID, JD) , therefore, $ID=I_1$; $JD=J_1+1$. Starting from the first neighbor label clockwise other neighbors of (I_1, J_1) as 1, 2, 3, 4, 5, 6, and 7. The coordinates of each points is shown in table 5.1. In this table $k_1=JD-J_1$ and $k_2=ID-I_1$.
3. The next move is controlled by the variable *dir* and it is determined by:
 - (a) $(dir+6) \bmod 8$ if *dir* is odd;
 - (b) $(dir+7) \bmod 8$ if *dir* is even;
4. The first pixel found with the same value as the current pixel is a new boundary element P_n .

Table 5.1: Co-ordinates of eight neighbors of a pixel

	LX(J)	LY(J)
1	ID	JD
2	LX(1)+k1	LY(1)-k2
3	LX(2)-k2	LY(2)-k1
4	LX(3)-k2	LY(3)-k1
5	LX(4)-k1	LY(4)+k2
6	LX(5)-k1	LY(5)+k2
7	LX(6)+k2	LY(6)+k1
8	LX(7)+k2	LY(7)+k1
if $ k1 = 1, k2 = 0$, if $ k2 = 1, k1 = 0$		

The *dir* value is updated accordingly.

5. Continue steps 3 and 4 and save the border points until the search reached at the other end of the capsule image.
6. Finally the detected inner border is represented by pixels P0.....Pn.

Since a major part of the time in border tracing is taken up in the computation of the coordinates of the neighboring pixels, fast calculation of the coordinates are very important specially in online applications. The usual approach to compute eight neighbors starting from the first neighbor is to use modulo arithmetic expressions. The simplest and fastest way to calculate the eight neighbors using only a single addition that is shown in table 5.1. The variables LX and LY are the (x, y) coordinates of the eight neighbors. It is also observed that the quantities k1 and k2 are either 1 in magnitude or 0.

5.5.4 Circle Fitting

The border points are extracted by tracing the boundary of the capsule's top and bottom part that was discussed in section 5.5.3. These points are then fitted to a circle using least square circle fitting and get an approximated circle. The basic equation for a circle is:

$$(x - x_c)^2 - (y - y_c)^2 = r^2 \quad (5.3)$$

where, (x_c, y_c) is the center and r is the radius and (x, y) are the border data points. In terms of parameters a, b, c, equation 5.3 can be written as:

$$x^2 + y^2 + ax + by + c = 0 \quad (5.4)$$

where, $a = -2x_c$ and $b = -2y_c$ and $c = x_c^2 + y_c^2 - r^2$

Now, using QR factorization (least square sense) the equation 5.4 is solved for the parameters a, b and c and then the center (x_c, y_c) and radius r of the circle is calculated using the following equations:

$$x_c = -\frac{a}{2} \quad (5.5)$$

$$y_c = -\frac{b}{2} \quad (5.6)$$

$$r = \sqrt{(x_c^2 + y_c^2) - c} \quad (5.7)$$

5.6 Simulation Results and Performance Evaluation

We have tested several image enhancement techniques using Matlab 7.0 in a Pentium IV, Intel Core 2 Duo @1.60GHz system. The image database of sample size more than 6000 was generated from the images captured from the developed capsule inspection system. The capsule size 0 was used in this application. The time requirement for inspecting one capsule from 4 different views is as follows. The industry requirement is to inspect 1000 capsule per minute. In the OptiSorter, 4 quadrants are working in parallel. Therefore, 250 capsules per min per quadrant have to be inspected. So, to inspect one capsule it requires $(60/250)=240$ ms. The hardware is developed in a way where around 50 ms is required for image acquisition. Therefore, around 190 ms is reserved for capsule inspection from 4 different views.

It is observed that bottom and top camera images carry more information than left and right camera images. That's why bottom and top camera images are used to measure almost all parameters like length, width, cap size, body size, closed joined length and number of pixels in each area. Most of the decisions are taken from these two images and if any parameters go out of range, it is marked as a defective capsule and sends it to the appropriate bin. Therefore, comparing the image size and parameters calculation, it is estimated that the top and bottom camera images require 60 ms each and right and left camera images require 35 ms each.

Our simulation results show that median filter provides best results both in terms of computational time and global variance. The enhanced image using different methods are shown in figure 5.9.

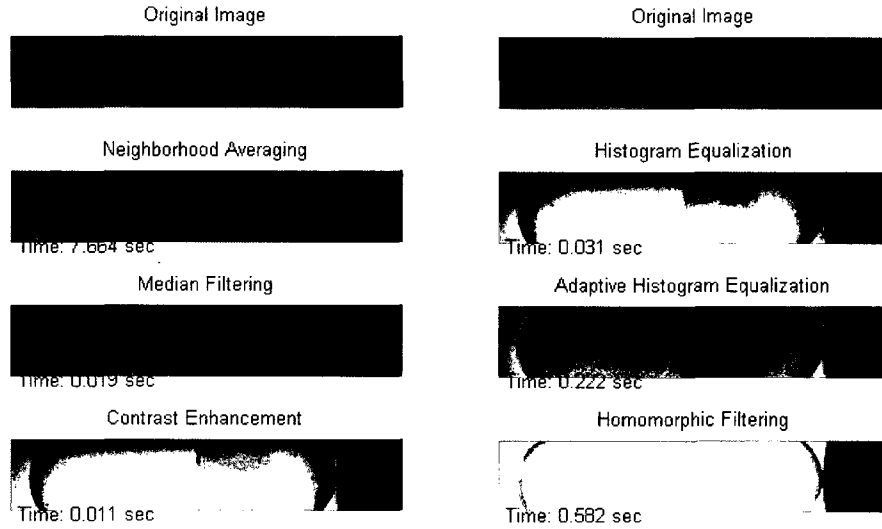


Figure 5.9: Image enhancement using different Techniques

The global variance and computational time is shown in figure 5.10. Comparing the enhancement quality in figure 5.9, global variance and computational time in figure 5.10, it is obvious that the median filter shows best performance in this application. The original image and enhanced image using median filtering is shown in figure 5.11. The enhanced image is then passed through the image segmentation phase.

The success of the capsule inspection system depends on the quality of the binary image. Different segmentation techniques were tested in this application. Otsu [37] and Kittler [38] are the dominating techniques in most survey that was discussed in the chapter 3. The binary image using these two methods is shown in figure 5.12.

ANN-based thresholding was investigated in this application and applied to binarize the capsule image using five dominant features according to the paper [67]. These five features pixel value, mean, standard deviation, smoothness and entropy described in the section 3.4 are calculated from the 5x5 neighborhood pixels [67] and used as input to the ANN to binarize the document image. In this application, the image features are also calculated from the 5x5 neighborhood pixels and save it to the training dataset as an object and background based on visual inspection. The training data preparation is shown in figure 5.13. The network is trained from the available training dataset and weights are calculated using the method described in section 3.5. The binary image using ANN-based thresholding method is shown in figure 5.14. The ANN architecture and computational time is shown in table 5.2.

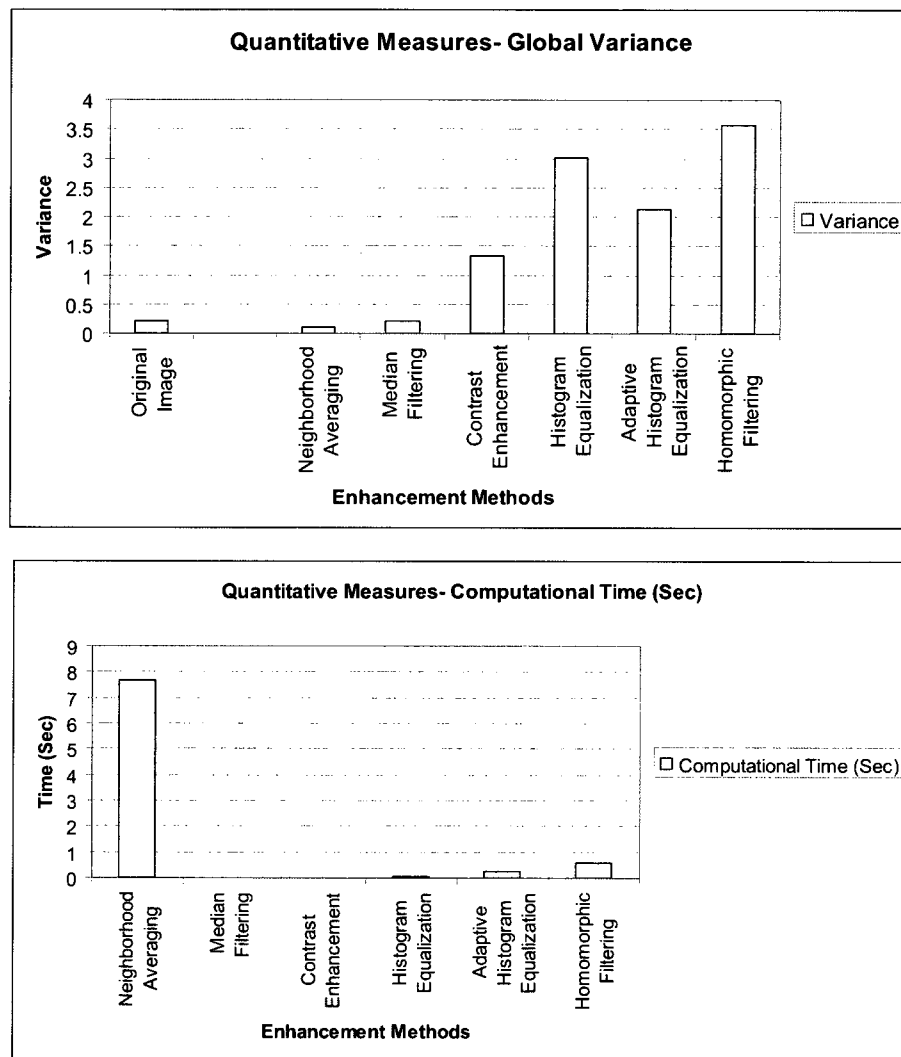


Figure 5.10: Quantitative measures of different image enhancement techniques

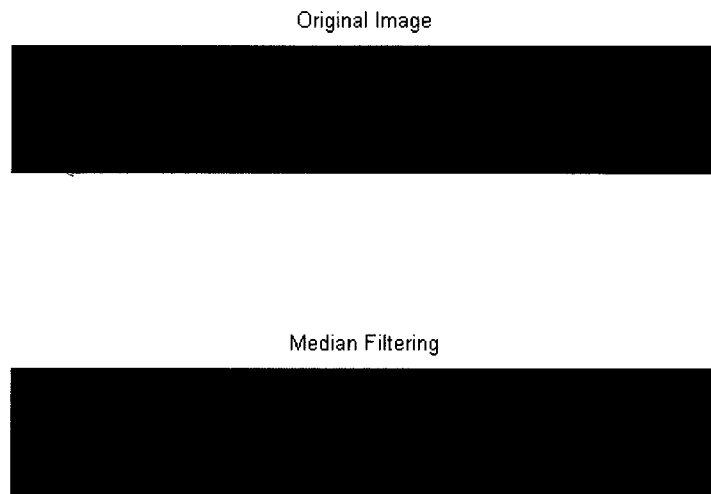


Figure 5.11: Image enhancement using Median filtering

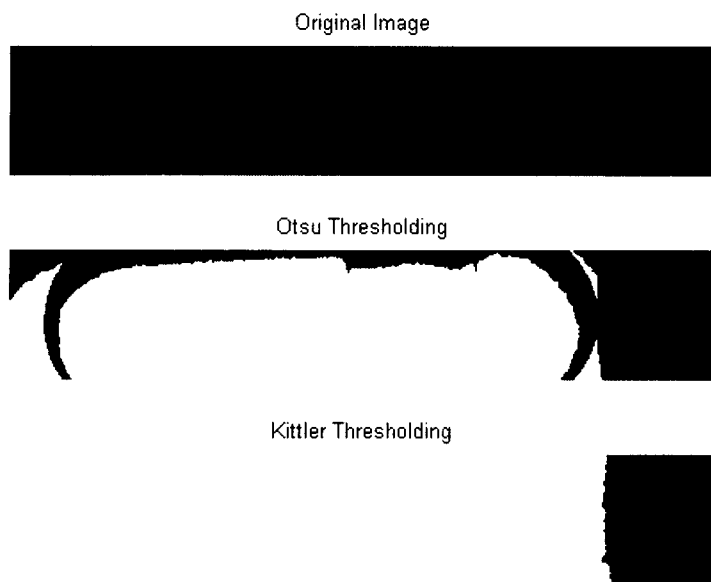


Figure 5.12: Binary image using Otsu [37] and Kittler [38]

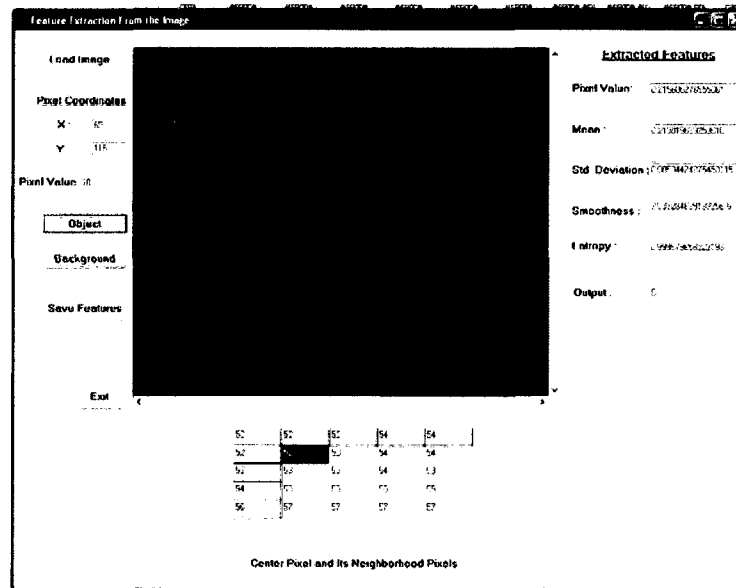


Figure 5.13: Training dataset generation using [67]

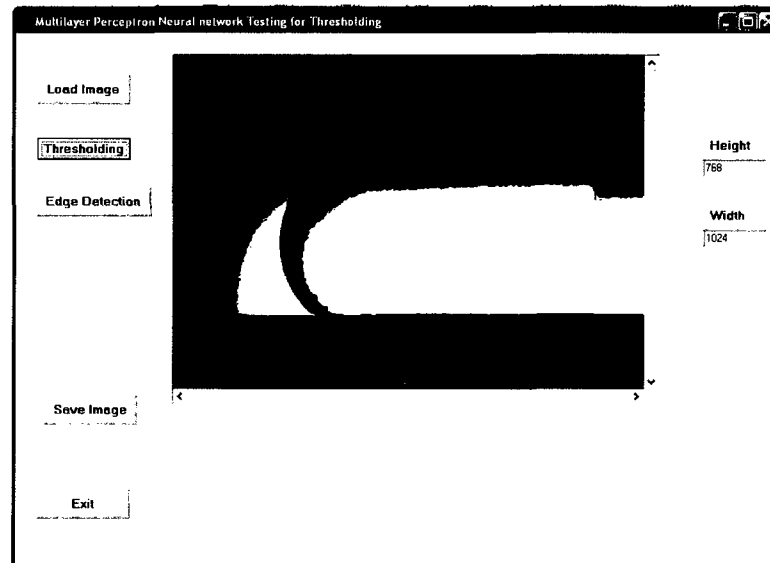


Figure 5.14: Binary image using ANN [67]

Table 5.2: Neural Network Architecture

Layer	No. of Nodes	Comments
Input	5	Features
Hidden	4	$2/3^*$ (Input+Output)
Output	1	1- Obj, 0-Back.
Time	2.54 sec (Computational)	80 ms (Expected)

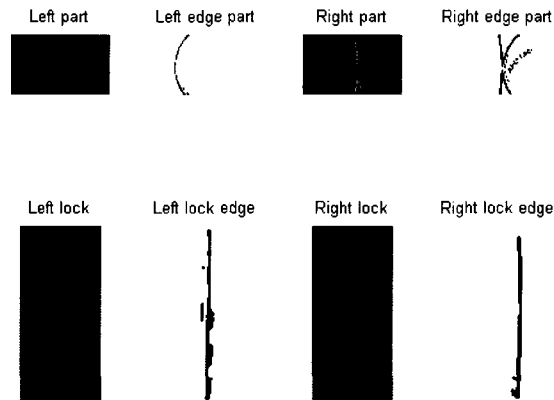


Figure 5.15: Edge-based segmentation

The ANN-based image segmentation is implemented using Borland C++. The processing time to have a binary image is around 2.54 sec whereas the time available for complete quality inspection including image segmentation is 60 ms. The quality of the binary image is evident from the figure 3.2. Therefore in terms of computational time and quality of the binary image ANN-based image segmentation is not suitable approach in this application.

On the other hand the binary image of different parts of the capsule using edge-based segmentation using Sobel edge detector is shown in figures 5.15 and 5.16 where the quality of the binary image is evident. Using the proposed method the sample good and dented capsule is shown in figures 5.17 and 5.18 respectively. The dented capsule in bottom part is shown in figure 5.19.

In case of left and right camera images, the same image processing techniques are applied like

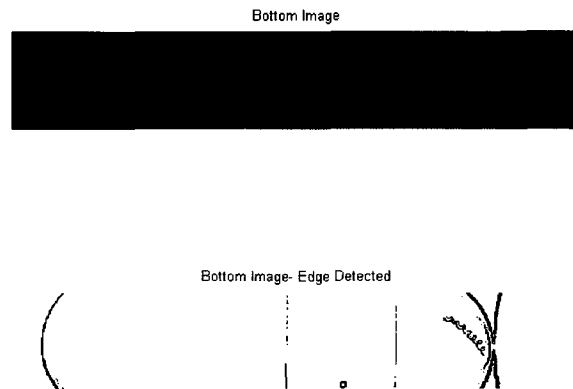


Figure 5.16: Closed view of edge-based segmentation

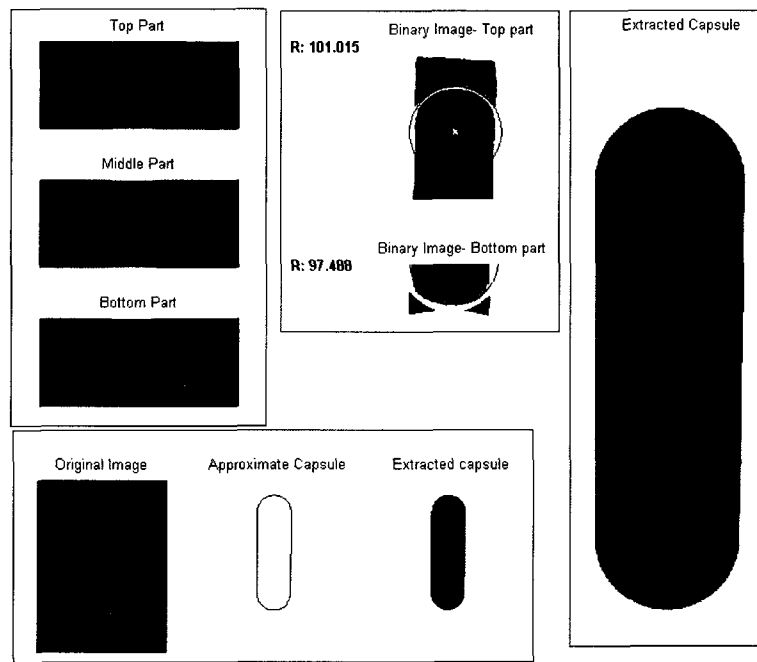


Figure 5.17: Good capsule image using the proposed method

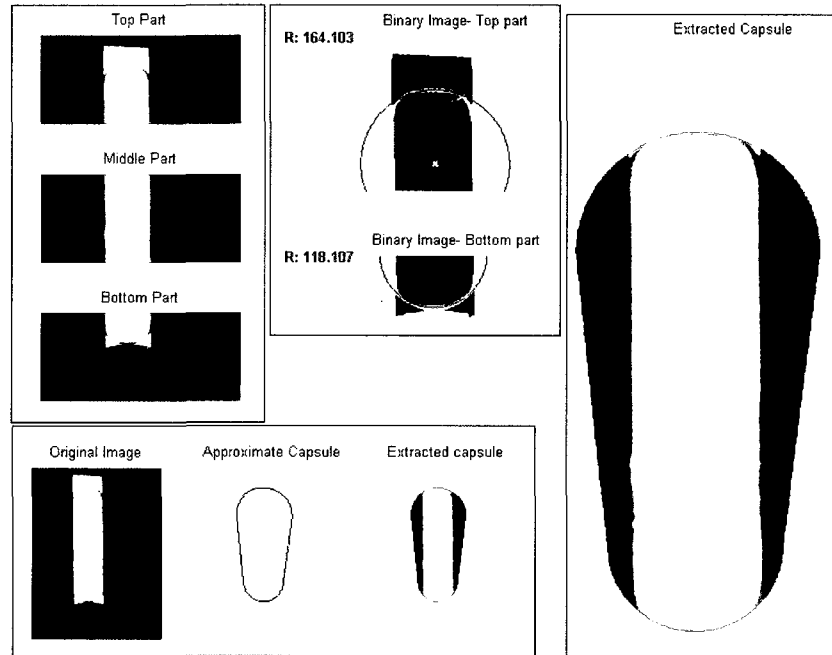


Figure 5.18: Dented capsule image using the proposed method- Example 1

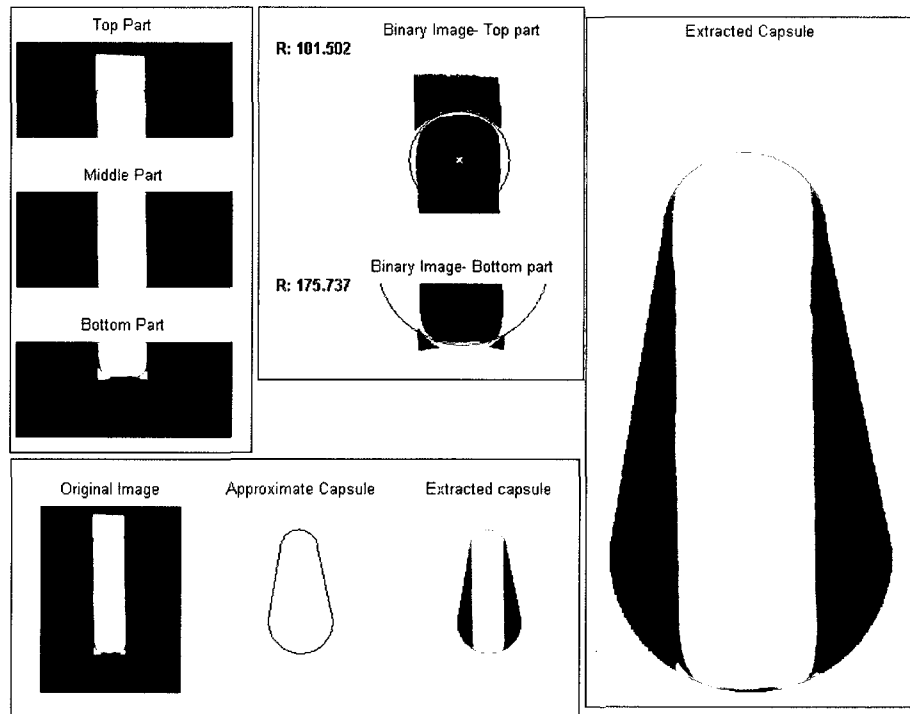


Figure 5.19: Dented capsule image using the proposed method- Example 2

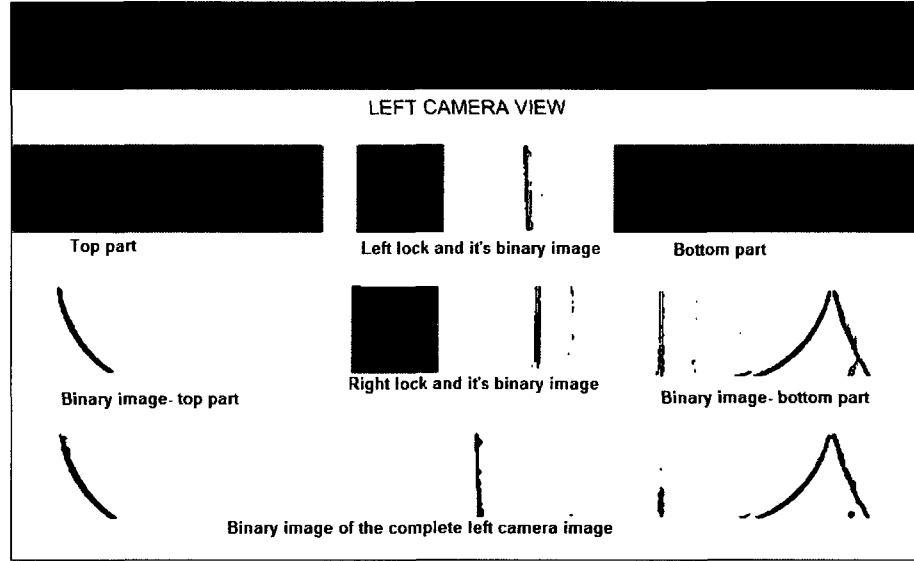


Figure 5.20: Binary image of left camera image using proposed method

median filtering, Sobel edge detector and Otsu thresholding [37] techniques. From the left and right camera images mostly the surface flaws are calculated because some flaws are not detected in bottom and top images and it can be detected from right and left images very clearly. But the image processing techniques are same as bottom and top images. In left and right images there is a similarity that makes the image processing task easier. The left image is the vertically flipped image of the right image. Therefore, the image processing techniques that are used to process the right image can be used to analyze the left image, only before processing the left image is vertically flipped. The binary image of different parts of the left and right camera images are shown in figure 5.20 and 5.21.

The complete quality inspection result using the edge-based proposed method is shown in table 5.3 for sample size approximately 6000 (size 0) where 4800 capsules are good and 1200 capsules are bad. Correct classification is obtained about 97.1% and missclassification is about 2.9%. The estimated sensitivity and specificity of the proposed system is shown in table 5.4. In this table it is clearly shown that the sensitivity of the proposed system is 97.45% and the specificity is 95.67% whereas the industry requirement was 95%. It is estimated that the computational time for all four images is around 190 ms and the computational time using proposed method requires approximately 160 ms.

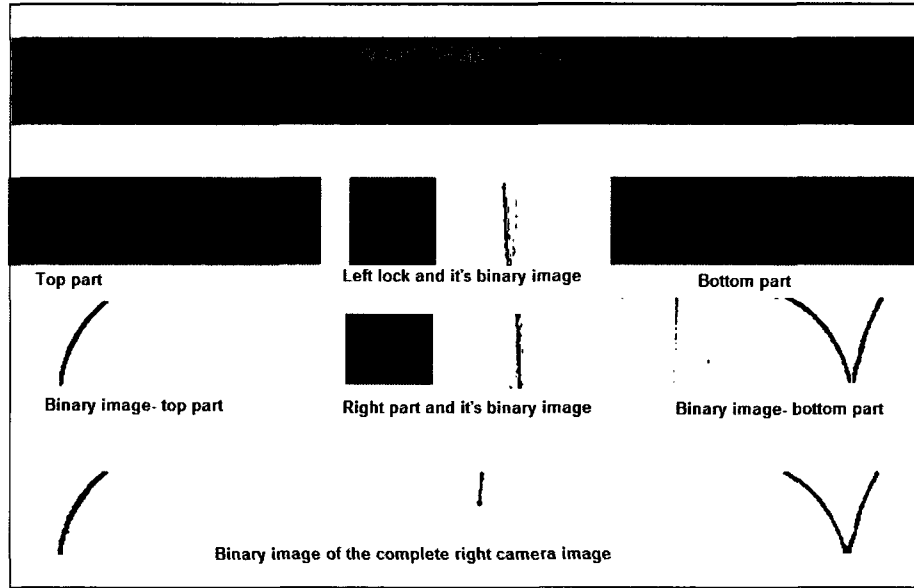


Figure 5.21: Binary image of right camera image using proposed method

5.6.1 Rationale Using Four Cameras

Three cameras are enough to cover full 360° , 120° apart from each other. However, it was found that for through inspection four cameras per quadrant provide better results. There are some bubbles, cracks or holes in a capsule that are missed when we use three cameras. This is because the orientation of the capsule is such that the defect is not clearly visible from bottom, left and right camera, hence it is missed in image segmentation stage. But if we use one more camera on top, it becomes clearly visible and the defect is detected easily. An example of a defective capsule with bubble is detected using top camera and is shown in figures 5.22 and 5.23.

5.7 Conclusion

In this chapter, we have presented an edge-based binarization method to extract capsule from the low contrast gray scale image. With the developed low cost hardware system, the image processing system meets the requirement of inspecting 1000 capsules per minute with acceptable accuracy of 95%. Dents are detected from the radius of the half circle generated from the border points of the top and bottom part of the capsule. The reference points for calculating the number of connected components are calculated from the left lock and right lock regions and two radius of top and bottom parts. The proposed system is flexible and a threshold can be set to detect holes, bubbles,

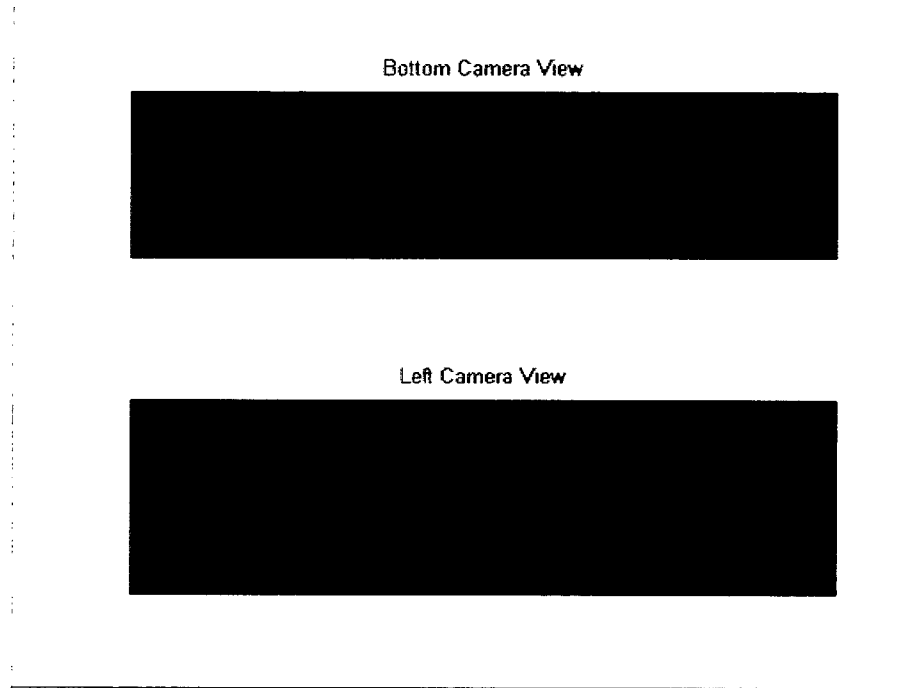


Figure 5.22: Bottom and left camera view of a defective capsule

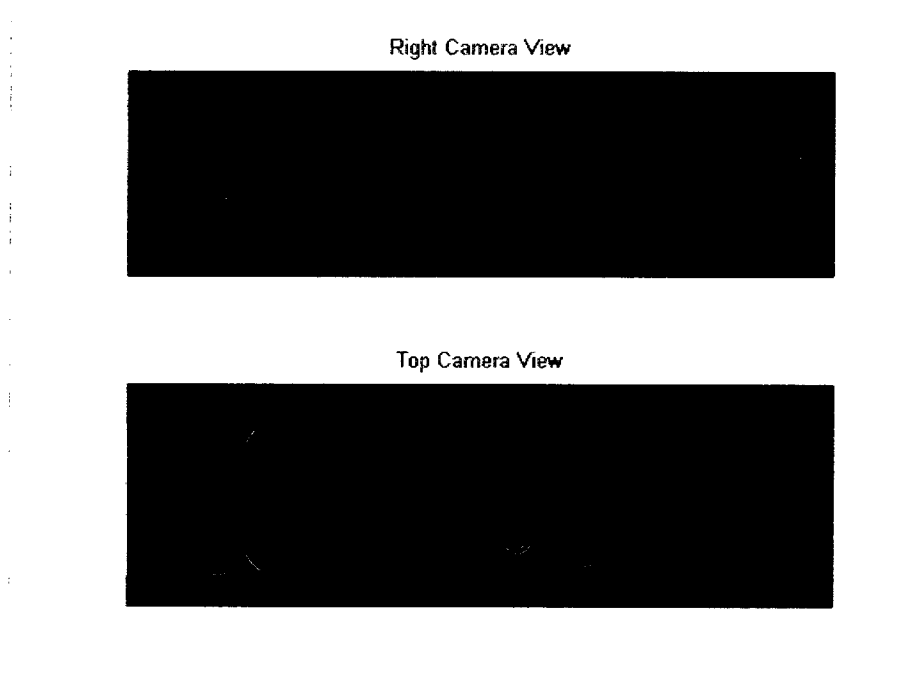


Figure 5.23: Right and top camera view of a defective capsule

Table 5.3: Quality Inspection Result

Typical Defects	False Positive (FP)	False Negative (FN)
Incorrect size	8	10
Dents	9	38
Cracks	17	25
Holes	10	23
Bubbles	8	26
Missing Cap	00	00
Closed Cap	00	00
Wrong Col	00	00
Total Misclassification	52	124
-	-	-
Misclassification (%)	2.9	
Correct Classification (%)	97.1	
Time Requirement	190 ms	
Computational Time (approx.)	160 ms	

cracks, dirt and any kind of foreign body inside the capsule of different size based on the industry requirement. Using the proposed system 97.1% correct recognition rate is achieved. The sensitivity of the proposed system is 97.45% and specificity is 95.67% which is very much promising. The proposed quality inspection of two part gelatin capsules not only maintain high throughput, but also performs accurately and reliably. It also meets cost and speed requirements and provides the flexibility to process different size capsules through its reconfigurable hardware and software.

Table 5.4: Performance Evaluation

True Positive (TP)- 4678	False Positive (FP)- 52
False Negative (FN)- 122	True Negative (TN)- 1148
Sensitivity= $TP/(TP+FN)$	97.47%
Specificity= $TN/(TN+FP)$	95.67%
Industry Requirement	95%

Chapter 6

ANN-based Mass Classification in Biomedical Imaging

6.1 Introduction

Breast cancer is the most common cause of cancer related deaths amongst women in the world. Early detection of breast cancer is the key to reduce the breast cancer mortality. X-ray mammography is currently considered the most popular, cost-effective, low radiation dose and relatively accurate method of early detection of breast carcinomas. Radiologists carefully search each mammogram for any visual sign of abnormality. However, retrospective studies shows that abnormalities are often embedded in and camouflaged by varying densities of breast tissue structures [80, 81] yielding the estimated sensitivity of the radiologists is about 75% during routine breast screening [1]. In order to increase the sensitivity, double reading can be one solution that is expensive. Another solution can be computer-aided-diagnosis (CAD) system that can be used as a second opinion to aid the radiologist by indicating the locations of suspicious abnormalities in mammograms based on their physical attributes. The final decision regarding the likelihood of the presence of a cancer depends on the radiologist. The principal stages of computer-aided breast cancer detection and classification is shown in figure 6.1.

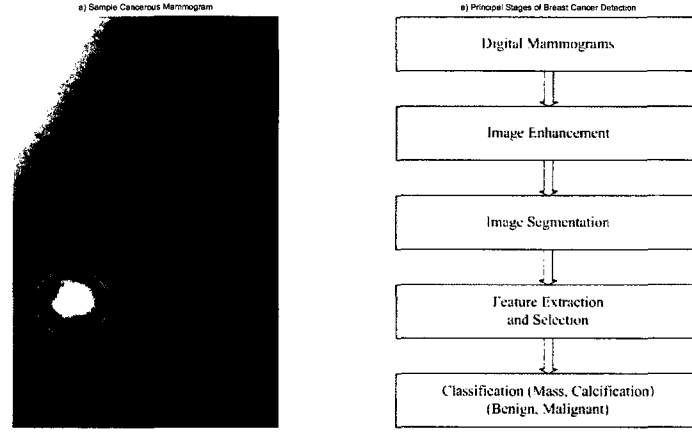


Figure 6.1: Sample mammogram and principal stages of breast cancer detection

6.2 Literature Survey

Masses and microcalcifications are two important early signs of breast cancer. Masses are often indistinguishable from the surrounding parenchyma because their features can be obscured or similar to the normal inhomogeneous breast tissues. The gray levels of those inhomogeneous tissues in the breast could vary with the distribution of breast soft tissue. Furthermore, the difficulty could be increased due to the fact that the masses in digitized mammograms are similar to the glands, cysts or dense portion of the breast [82]. This makes the automatic mass detection and classification challenging. The main aim of this thesis is to develop a CAD system for breast cancer diagnosis and detection based on automated segmentation of masses in mammograms.

In recent years, a few researchers in either academia or industry have used different approaches to do the segmentation and classification of masses. The segmentation and classification steps are vital for the performance of the CAD system that is shown in figure 6.1. Many involve the use of segmentation techniques to separate the mass called region of interest (ROI) from the breast tissue. Examples of these segmentation techniques include:

Histogram-based methods [83, 84, 85] are classified into two categories, global thresholding and local thresholding. Global thresholding [83, 84] methods are based on global information, such as histogram of the mammograms. It is widely used and easy to implement. But it is not good for identifying ROIs and false-positive (FP) and false-negatives (FN) may be too high. In local thresholding [86, 87], the threshold value is determined locally. It can refine the results of global

thresholding, and is better for mass detection than global thresholding. It can not accurately separate the pixels into suitable sets. It is often used as an initialization of other algorithms. In template-matching [88, 89, 90]- the possible masses are segmented from the background using prototypes. It is easy to implement. If the prototypes are appropriate, it can provide good results. It depends on the prior information of the masses, it may result high number of FPs. The region growing [91, 87, 92] methods find a set of seed pixels first, then grow iteratively and aggregate with the pixels that have similar properties. In these methods the main challenge is to find the seed points because of the peculiarity of these kind of object of interest. Bilateral image subtraction [93] is based on the normal symmetry between the left and right breast. It is easy to implement, and the difference between the left and right mammogram images can be identified as suspicious regions but it is difficult to register the left and right breast correctly. The fuzzy techniques [94] including fuzzy thresholding and fuzzy region growing. It can handle the unclear boundary between normal tissue and tumors but it is not easy to determine suitable membership functions and rules. Other methods include edge-based methods [85, 95] that detects the discontinuity in mammograms.

Features extraction and selection is a key step in mass detection and classification. Features are calculated from the region of interest (ROI) characteristics such as size, shape, density and smoothness etc. [96]. Feature space is very large and complex due to wide diversity of the normal tissues and the variety of the abnormalities. Feature space can be divided into 3 sub-spaces [97] intensity features, shape features and texture features. Haralick texture features [98] are based on the gray-level co-occurrence matrix (GLCM) or gray-level dependence matrix (GLCD) and those features are used for seed selection for mass segmentation. It will be discussed in details in section 6.3.

Once the features are extracted and selected, the features are input into a classifier to classify the detected suspicious areas into benign or malignant. Classifiers such as linear discriminants analysis (LDA) and ANN have performed well in mass classification.

LDA's [99, 100] are traditional method for classification. They construct decision boundaries by optimizing certain criteria to classify cases into one of mutually exclusive classes. They show high performance for linear separable problems but poor for non linear separable data.

ANN [101, 102] are invaluable tools in various medical diagnostic systems. Lisboa [103] reviewed the improvements in healthcare arising from the participation of ANN in medical field. The key attributes like distributed representations, local operations, and non-linear processing make ANN suitable for taking few difficult decisions from massive amount of data. Thus when expert knowledge is unavailable in full-fledged sense as for example in case of masses, ANN provides alternative and

better solutions. It is robust, no rule or explicit expression is needed and widely applicable [104]. But there is no common rule to determine the size of the ANNs, long training time and sometimes over training. In this thesis, we will propose an automated seeded region growing for mass extraction from the breast tissue using Haralick texture features called ROI and then classify them using ANN classifier based on statistical and textural features.

6.3 Haralick Texture Features [98, 105]

The Haralick texture features are used for image classification. The basis for these features is the Gray-Level Co-occurrence Matrix (GLCM) or Spatial Gray Level Dependence (SGLD) Matrix. The calculation of Haralick texture feature can be broken down into two parts or modules;

1. The construction of the GLCM and
2. The calculation of the 13 texture features based on the co-occurrence matrices.

Some of these features include angular second moment, contrast, correlation, as well as a variety of entropy measures.

6.3.1 Gray Level Co-occurrence Matrix

GLCM is a square matrix with dimension N_g shown in matrix 6.1, where N_g is the number of gray levels in the image. It is used to describe the patterns of neighboring pixels in an image at a given distance, d . Each element $[i, j]$ of the matrix is generated by counting the number of times a pixel with value i is adjacent to a pixel with value j .

$$\begin{bmatrix} P(1, 1) & P(1, 2)..... & P(1, N_g) \\ P(2, 1) & P(2, 2)..... & P(2, N_g) \\ P(3, 1) & P(3, 2)..... & P(3, N_g) \\ \dots & \dots\dots\dots & \dots \\ P(N_g, 1) & P(N_g, 2)..... & P(N_g, N_g) \end{bmatrix} \quad (6.1)$$

To calculate the texture features, 4 matrices are needed to describe different orientations. Because, adjacency can be defined to occur in four directions in a 2D square pixel image. They are, 1) Horizontal direction (0°)- describes pixels that are adjacent to one another horizontally in 0° direction.

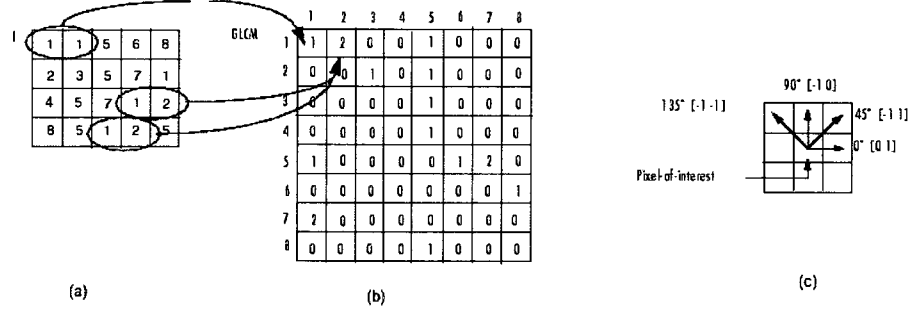


Figure 6.2: Example of GLCM for $d = 1$: (a) The original image (b) Final result of the horizontal GLCM $d = 1$. c) Four directions (0° , 45° , 90° , 135°)

- 2) Vertical direction (90°)- describes pixels that are adjacent to one another vertically in 90° direction.
- 3) Diagonal directions (45° and 135°)- describes pixels that are adjacent to one another both in left and right diagonal directions.

Figure 6.2 gives a graphical description of this process for horizontal direction. The method to apply this to the other 3 orientations is done in a similar fashion. Rotation invariance is a primary criterion for any features used with these images; a kind of invariance was achieved for each of these statistics by averaging them over the four directional co-occurrence matrices.

6.3.2 Statistical Properties of GLCM and Haralick Features Calculation

Some statistical features can be derived from the GLCM that are needed to calculate 13 Haralick texture features. Statistical features are shown in equation 6.2 to 6.9.

$$R = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} P(i, j) \quad (6.2)$$

$$p(i, j) = \frac{P(i, j)}{R} \quad (6.3)$$

$$p_x(i) = \sum_{j=1}^{N_g} p(i, j) \quad (6.4)$$

$$p_y(j) = \sum_{i=1}^{N_g} p(i, j) \quad (6.5)$$

$$p_{x+y}(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) \quad (6.6)$$

where, $i + j = k$ and $k = 2, 3, 4, \dots, 2N_g$ and x and y are the co-ordinates (row and column) of an entry in the co-occurrence matrix and $p_{x+y}(k)$ is the probability of GLCM coordinates summing to $x + y$.

$$p_{x-y}(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) \quad (6.7)$$

where, $|i - j| = k$ and $k = 0, 1, 2, \dots, N_g - 1$

$$HXY1 = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) \log(p_x(i)p_y(j)) \quad (6.8)$$

$$HXY2 = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p_x(i)p_y(j) \log(p_x(i)p_y(j)) \quad (6.9)$$

Haralick texture features are calculated from these above statistical properties of GLCM. These features are shown in equation 6.10 to 6.22.

1. Angular second moment

$$H1 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j)^2 \quad (6.10)$$

2. Contrast

$$H2 = \sum_{k=0}^{N_g-1} k^2 p_{x-y}(k) \quad (6.11)$$

3. Correlation

$$H3 = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} ijp(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (6.12)$$

where μ_x , μ_y , σ_x and σ_y are the means and standard deviations of p_x and p_y , the partial probability density functions.

4. Sum of squares: Variance

$$H4 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - \mu)^2 p(i, j) \quad (6.13)$$

5. Inverse difference moment

$$H5 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{1}{1 + (i - j)^2} p(i, j) \quad (6.14)$$

6. Sum average

$$H6 = \sum_{i=2}^{2N_g} ip_{x+y}(i) \quad (6.15)$$

7. Sum entropy

$$H7 = - \sum_{i=2}^{2N_g} ip_{x+y}(i) \log(p_{x+y}(i)) \quad (6.16)$$

8. Sum variance

$$H8 = \sum_{i=2}^{2N_g} (i - H7)^2 p_{x+y}(i) \quad (6.17)$$

9. Entropy

$$H9 = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) \log(p(i, j)) \quad (6.18)$$

10. Difference variance

$$H10 = \sum_{i=0}^{N_g-1} i^2 p_{x-y}(i) \quad (6.19)$$

11. Difference entropy

$$H11 = - \sum_{i=0}^{N_g-1} p_{x-y}(i) \log(p_{x-y}(i)) \quad (6.20)$$

12. Information measure of correlation 1

$$H12 = \frac{H9 - HXY1}{\max(HX, HY)} \quad (6.21)$$

HX and HY are the entropies of p_x and p_y

13. Information measure of correlation 2

$$H13 = \sqrt{1 - \exp^{-2(HXY2 - H9)}} \quad (6.22)$$

6.4 Proposed Methods

In this section, we propose an automatic seeded region growing method to extract the mass from the ROI. Haralick texture features are used to select the seed from the ROI from where the region growing starts and then grow iteratively and aggregate with the pixels that have similar properties. Then the statistical and textural features described in section 3.4 mean, standard deviation, entropy, smoothness, skewness, kurtosis and uniformity are extracted from the extracted mass to use them for classification into benign or malignant. The different steps to extract the mass and classification are as follows:

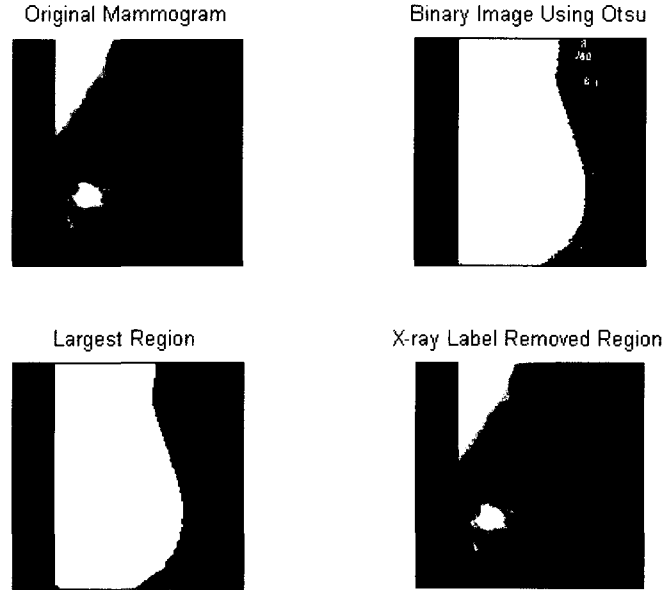


Figure 6.3: X-ray label removal from the original mammogram

6.4.1 ROI Preprocessing

The size of a digitized mammogram is generally very large but the ROIs are very small and limited to areas being determined as suspicious regions of masses such as shown in figure 6.1. It takes long time to process the background although we are interested in ROI only. So, the first step is to separate the ROIs from the image background so that the image processing will not be overwhelmed and dominated by the large image background. To do that x-ray removal, pectoral muscle removal, breast region and ROI extraction are the key steps.

a) X-ray label removal

To remove the x-ray label, a global histogram is calculated and then Otsu [37] thresholding technique is applied to have a binary image. The result of applying the thresholding technique is a collection of different homogeneous regions in the image. Using Connected Component Labeling [31, 106] the largest region is extracted which is the union of breast and pectoral muscle. Figure 6.3 shows the steps followed from the original mammogram to obtain the x-ray label removed region.

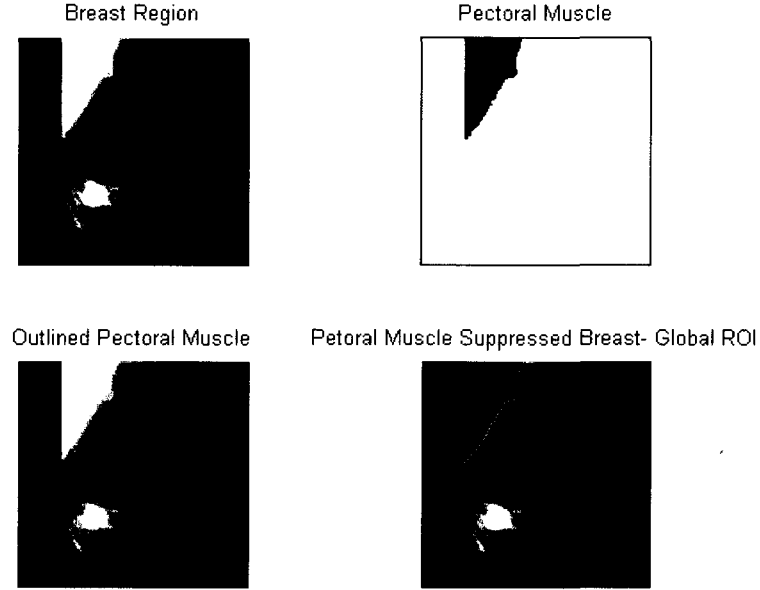


Figure 6.4: Breast extraction- Global ROI

b) Pectoral muscle removal- Global ROI

Mass and pectoral region may have similar texture characteristics, causing a high number of FPs when detecting suspicious masses [107]. It is slightly brighter compared to the rest of the breast tissue. In mediolateral oblique (MLO) mammogram it appears as a triangular region across the upper posterior margin of the image. To remove the pectoral muscle, a simple region growing technique is applied to the x-ray label removed region. First pixel on the non-curved side of the breast region from top-left side is considered as a seed pixel. Region growing starts from this seed and grow iteratively to fill the whole region of the pectoral muscle. Finally, the pectoral muscle is suppressed from the breast region and morphological operators are used to smooth the boundary. Figure 6.4 shows the steps followed from the original mammogram to obtain the pectoral muscle suppressed breast called global ROI.

c) ROI extraction

In order to extract suspicious regions (ROI) from the whole mammographic image a preliminary automatic identification is done that probably contain massive lesions based on the histogram analysis [83] of the global ROI. Mass is slightly brighter than its surrounding areas, produces a sharp peak

of unusual gray level intensity pixels. To extract the suspicious region, histogram is calculated of the global ROI. Then the peak of the histogram is obtained. To determine the existence of a peak, its width must be not more than 4 gray levels and it's height must be atypical compared with the height of the gray level of his environment. The steps for ROI extraction [83] are as follows:

1. Localize all points I (gray level) in the histogram that constitutes a peak. They must satisfy two conditions $h(I) > h(I - 1)$ and $h(I) > h(I + 1)$, where $h(I)$ represents the number of pixels in the image that have the gray level intensity I .

2. The height of the peak I are as follows:

$$H_I = \frac{H_L + H_R}{2} \quad (6.23)$$

H_L and H_R are the left height and right height respectively that can be defined as:

$$H_L = \max \{h(I) - h(I - 1), h(I) - h(I - 2)\} \quad (6.24)$$

$$H_R = \max \{h(I) - h(I + 1), h(I) - h(I + 2)\} \quad (6.25)$$

3. Each H_I detected in the previous step in the histogram is divided by the average of the height of the base of the peak using the equation 6.26

$$R_I = H_I / \frac{h(I_L) + h(I_R)}{2} \quad (6.26)$$

where $h(I_L)$ and $h(I_R)$ corresponds with the values of the base of the peak on its left and right side respectively.

4. All the gray levels in the histogram other than peak are replaced by 0 and all peaks $h(I)$ are replaced by R_I .

5. Finally all values R_I smaller than a determined threshold U are eliminated. The threshold value in this application was found 0.35 from the experiment conducted by us. If it is less than that value, the number of ROIs becomes more where most of them are normal breast tissue. If it greater than that, system may miss some ROIs. From the experiment this value was found to be optimal.

6. For each one of the peaks I detected in the previous step will be segmented from the original image where all pixels with gray level value $I(i, j)$ in $I \pm \epsilon$ will be replaced by white color following the condition below for all $\forall(i, j)$, if

$$I(i, j) \in [I - \epsilon, I + \epsilon] \rightarrow I(i, j) = 255 \quad (6.27)$$

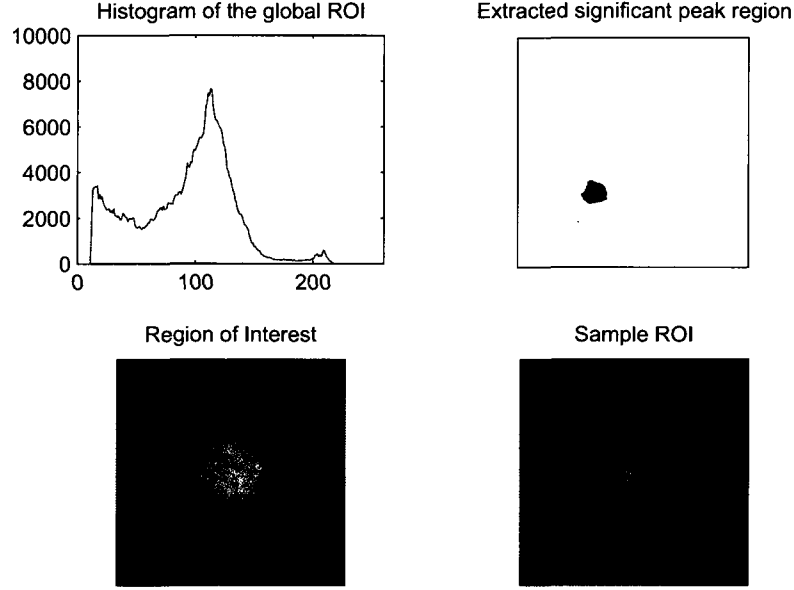


Figure 6.5: ROI extraction and sample ROI

for practical cases $\epsilon = 5$ is obtained from experiment.

7. Each of the segmented regions that is found to be suspicious region called ROI. The ROI is extracted from the original image keeping the suspicious region at the center of the image. In this application, ROI size is set at 256x256 whereas the original image size is 1024x1024.

Following the above procedure the sample ROI is shown in figure 6.5.

6.4.2 Mass Extraction

Once the ROI is extracted, the next part is to extract mass from the ROI. The proposed system consists of contrast enhancement, segmentation by automatic seeded region growing algorithm using Haralick texture features.

a) Contrast Enhancement

The contrast of the ROI is enhanced using the following nonlinear operator [91].

$$I_{en}(i, j) = \left(\frac{I(i, j)}{I_{max}} \right)^k * I_{max} \quad (6.28)$$

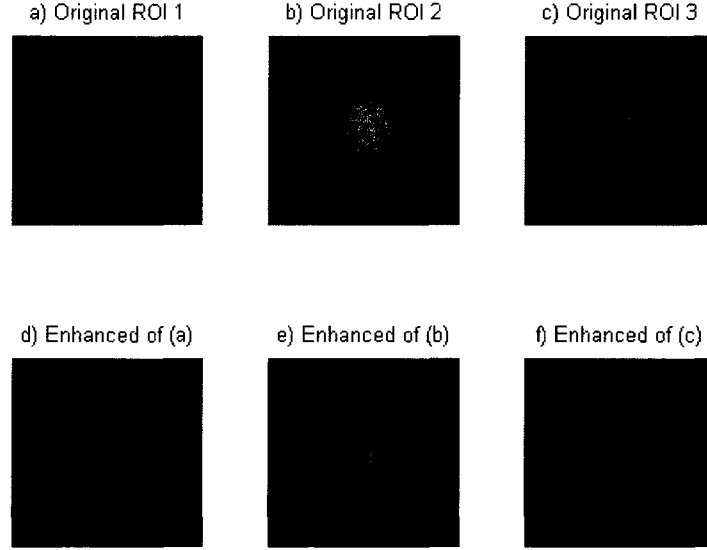


Figure 6.6: Sample Original and Enhanced ROIs

where $k = 2, 3, 4, \dots$ and $I(i, j)$, $I_{en}(i, j)$ are the pixel intensity of the original image and enhanced image respectively. I_{max} is the maximum intensity of the original image. This way we penalize the dark pixels more than the bright pixels [91]. The original ROI and enhanced ROI is shown in figure 6.6.

b) Automatic Seeded Region Growing Using Haralick Texture Features

Region growing approaches exploit the important fact that pixels which are close together have similar gray levels. The process starts from one or more points called seed points [91]. Then region growing expands the selected areas around the seeds to include nearby pixels falling within a threshold. In this algorithm the main challenge is to select the seeds. We propose an automated method to select the seeds. The steps are as follows.

1. The first step is to divide enhanced ROI into $R \times R$ non-overlapping blocks. Block size is application dependent and is very important factor in this algorithm. Because if block is too small, the difference of the mass textures from normal textures can not be well characterized. If it is too large, the result may be too coarse. In this application 32×32 block size was found satisfactory. Figure 6.7 shows the sample blocks division and the seed block location.

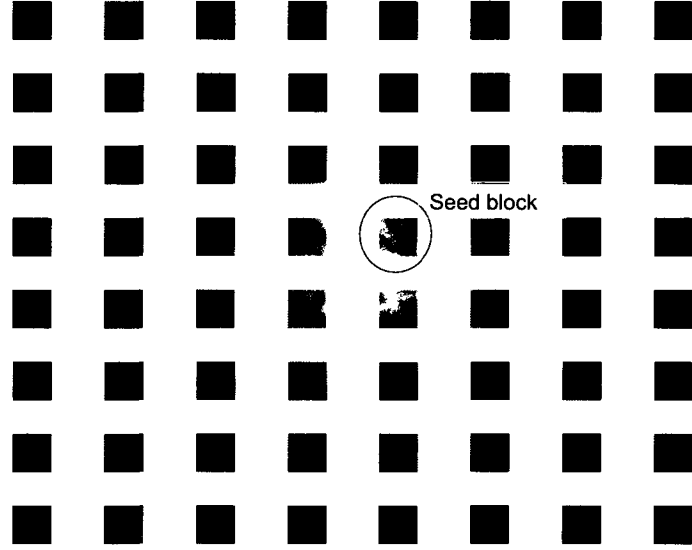


Figure 6.7: ROI division into blocks and seed block detection

2. Calculate the Haralick texture features from Spatial Gray Level Dependence Matrix (SGLD) of each block. Then select the significant features that can easily discriminate mass and non mass region. The normalized output of the Haralick texture features are shown in figures 6.8, 6.9 and 6.10 for an enhanced ROI of figure 6.6(d). It was observed that Out of 13 Haralick texture features Sum Average is found to be very significant to discriminate mass and non mass blocks that is described in equation 6.15.

3. Select the block that contains mass based on the feature selected in step 3. Maximum gray level of that block is proposed as the seed point.

4. Region growing starts from that point and then grow iteratively and aggregate with the pixels that have similar properties results in segmented mass region. The segmented image is smoothed using morphological operators.

5. Extract the contour of the segmented mass and approximate it to a circle. Estimate the radius of the circle and compare it with the ground-truth data. This comparison will provide the results how close the segmented mass to the ground-truth mass determined by the expert radiologists.

6. Extract the mass region from the original image that is used as an input for classification. Figure 6.11 shows the flowchart of the proposed automatic seeded region growing method and that

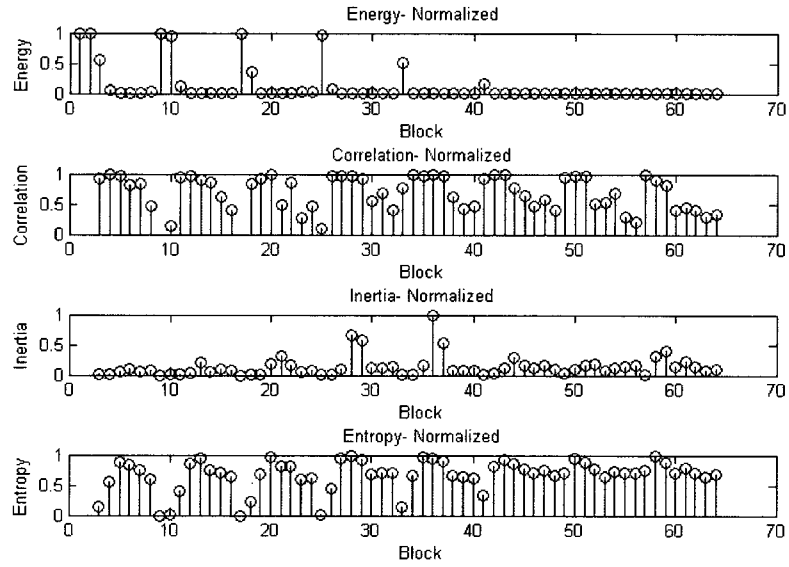


Figure 6.8: Haralick Texture Features (1-4)

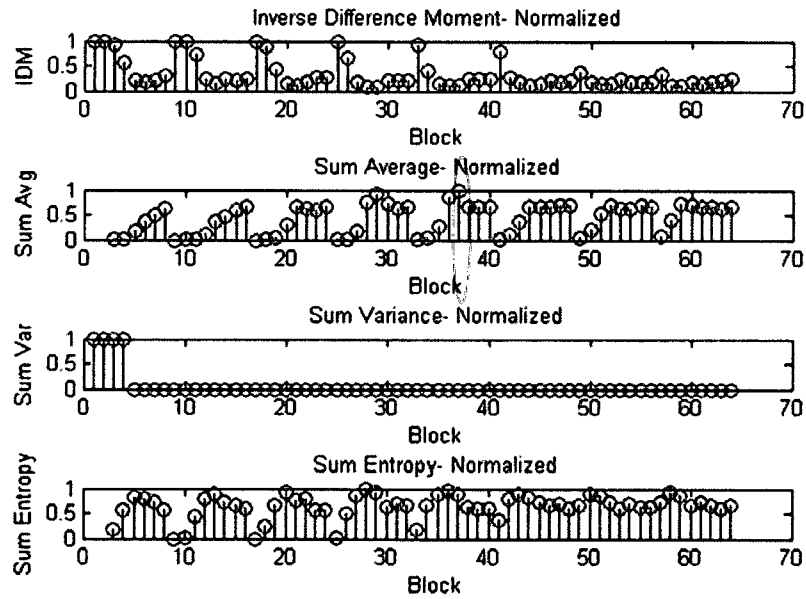


Figure 6.9: Haralick Texture Features (5-8)

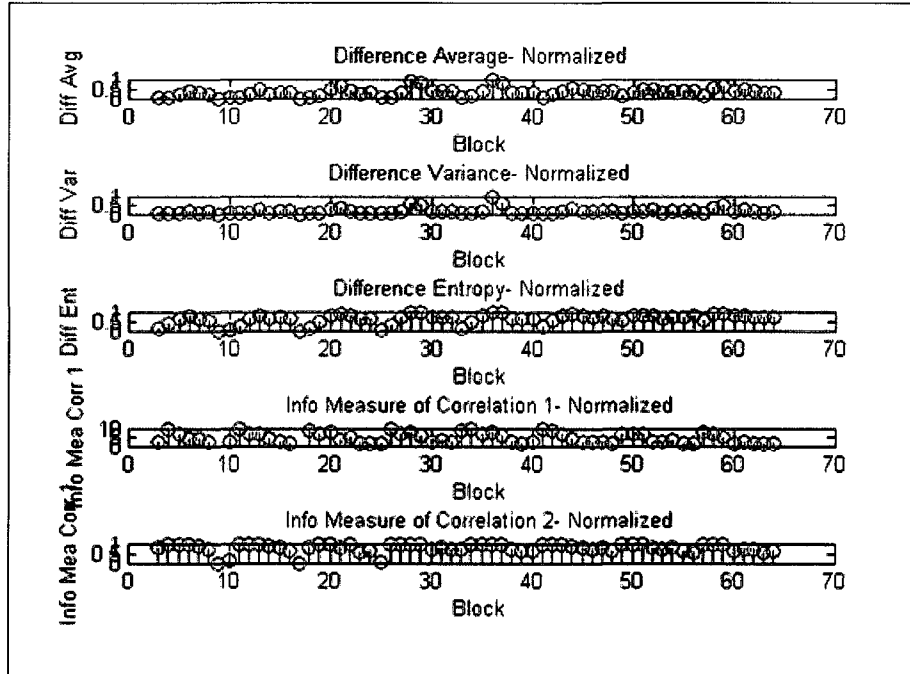


Figure 6.10: Haralick Texture Features (9-13)

mass using this method is shown in figure 6.12.

6.4.3 Mass Classification

One of the major mammographic characteristics for mass detection and classification is texture. ANN exploits this important factor to classify the mass into benign or malignant. The statistical textural features used in characterizing the masses are mean, standard deviation, entropy, skewness, kurtosis and uniformity [63]. These 7 features are used in preparing the training data for multi-layer perceptron (MLP) neural network which are obtained from the whole extracted mass region. The calculated 7 features and their corresponding target value (for benign=0 and malignant=1) are stored in a file and then used as inputs to the ANN to train the network to produce the weights needed for testing the classifier. Figure 6.13 shows the sample screen capture of training data preparation and classification.

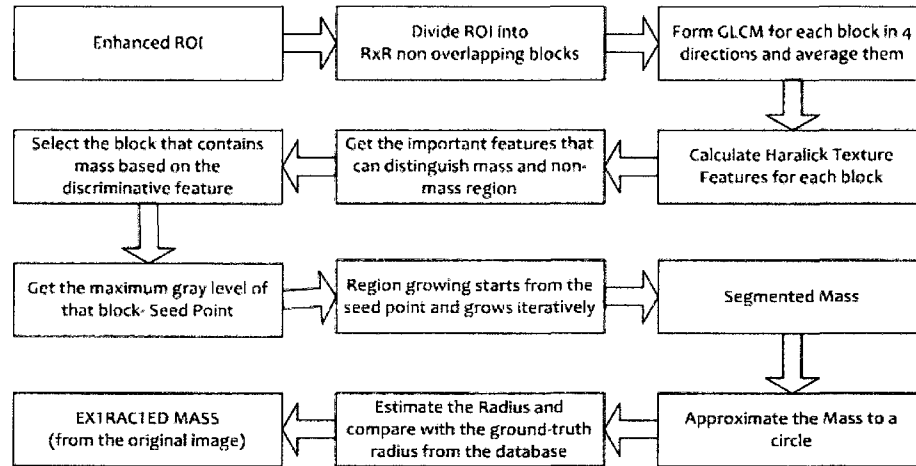


Figure 6.11: Flow chart of the proposed automatic seeded region growing method

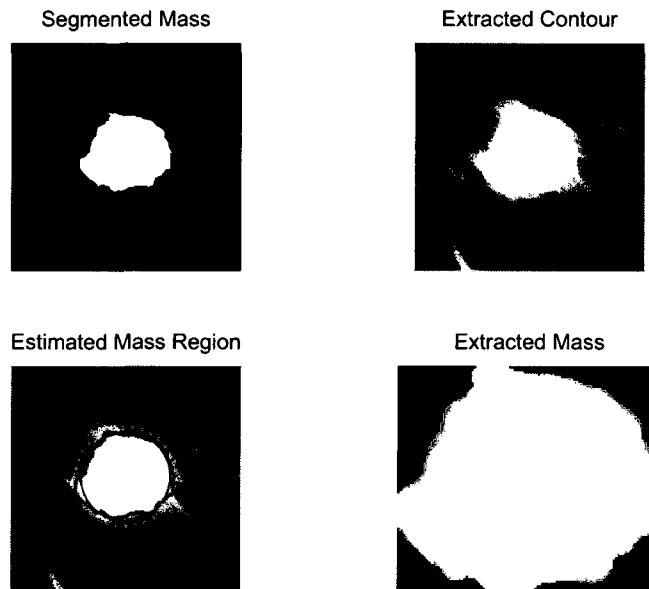


Figure 6.12: Extracted mass from the ROI

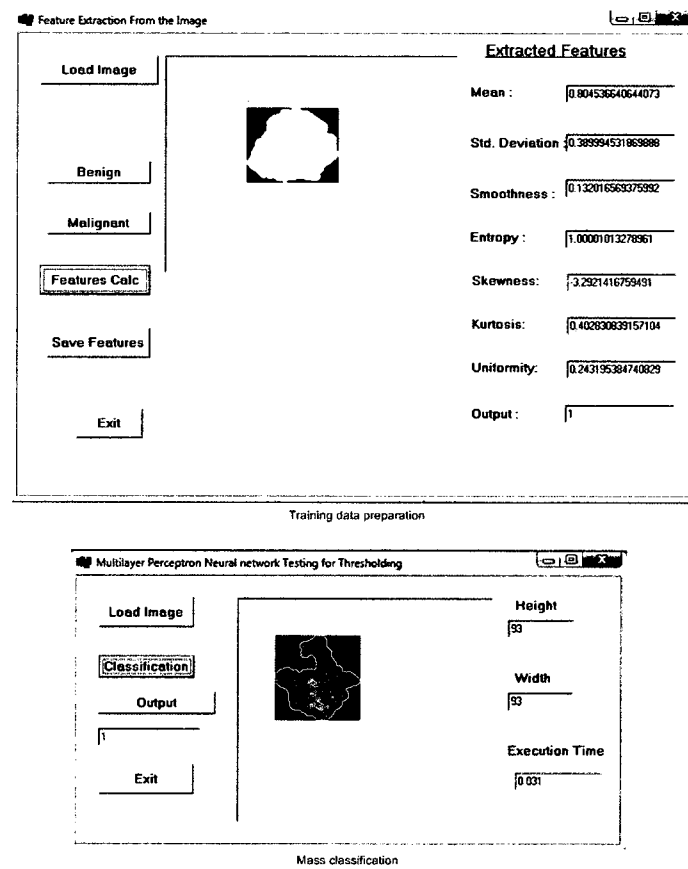


Figure 6.13: Training data preparation and mass classification

Table 6.1: Simulation Result: Mass Segmentation

Correct Segmentation (%)	Radiologist Sensitivity (%)	Incorrect Segmentation (%)
84.15	75	15.85

6.5 Simulation Results and Performance Evaluation

6.5.1 Image Database

To develop and evaluate the proposed system we used the Mammographic Image Analysis Society (MiniMIAS) [108] database. It is an organization of UK research group. Films were taken from UK National Breast Screening Programme that includes radiologist's "truth"-markings on the locations of any abnormalities that may be present. Images are available online at the Pilot European Image Processing Archive (PEIPA) at the University of Essex. This database contains left and right breast images for a total of 161 (322 images) patients with ages between 50 and 65. All images are digitized at a resolution of 1024x1024 pixels and at 8-bit grey scale level. The existing data in the collection consists of the location of the abnormality (like the center of a circle surrounding the tumor), its radius, breast position (left or right), type of breast tissues (fatty, fatty-glandular and dense) and tumor type if it exists (benign or malign). Each of the abnormalities has been diagnosed and confirmed by a biopsy to indicate its severity: benign or malignant. In this database, 42 images contain abnormalities (malignant masses) and 212 images are classed as normal and rest of them either contains microcalcifications or benign.

6.5.2 Results and Performance

Among the region-based approaches [91], the region growing algorithm appears as the natural choice in the mass segmentation, since the peculiarity of these kind of objects of interest is the connectivity of pixels, neither edge or luminance alone be used for the isolation of the region inside the mass. The proposed seed selection method made the process easier as seed selection is the major challenge in the region growing approach. This proposed method is applied to 82 benign and malignant images in MIAS database and the results obtained is shown in table 6.1. Samples masses both malignant and benign and extracted contour of the masses in the ROI is shown in figure 6.14 and 6.15 respectively.

The performance of the proposed algorithm is assessed by comparing the segmented area by our algorithm with, as a ground truth, the area within a radiologist marked contour [91]. The terms that are used for that purpose are:

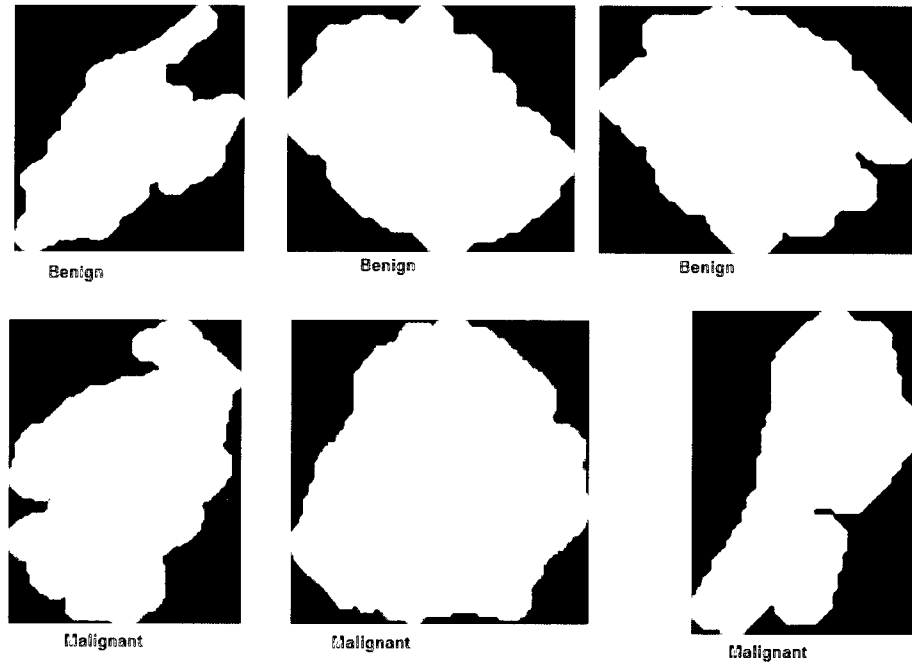


Figure 6.14: Sample extracted mass using proposed method

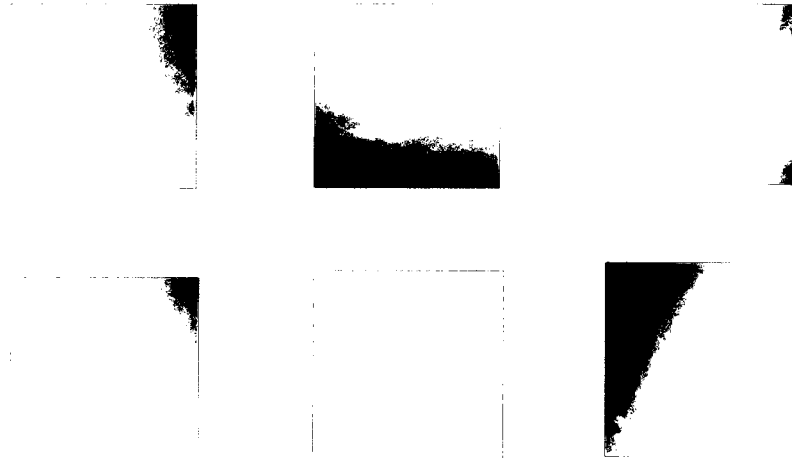


Figure 6.15: Sample extracted contour using proposed method

1. Estimated region (ER)- is the segmented region by our algorithm,
2. Reference region (RR)- circular area estimated by the radiologist,
3. Area difference (AD)- difference between RR and ER,
4. True positive (TP)- intersection of ER and RR,
5. False positive (FP)- the area not identified in RR,
6. False negative (FN)- the area in RR not identified in ER

and completeness (CM) and correctness (CR) are defined as follows:

$$CM = \frac{TP}{TP + FN}$$
$$CR = \frac{TP}{TP + FP}$$

In particular completeness is the percentage of the ground-truth region which is detected by the segmented region, while correctness is the percentage of correctly extracted breast region. The average completeness and correctness we obtained for 5 images are 78% and 94% respectively. It is observed that the completeness is below 80% for 5 segmented images. This is because the mass is estimated as a circle which is not the case in practice. Figure 6.16 and 6.17 shows the different regions that were obtained for calculating for CM and CR. For figure 6.16 the obtained CM and CR are 75% and 98% respectively and the CM and CR of figure 6.17 are 87% and 95% respectively.

ANN is used to classify the extracted mass into benign or malignant. The table 6.2 shows the proposed ANN structure. The proposed structure has 3 layers comprised of 7 units in input layer, 5 units in hidden layer and 1 unit in the output layer. 5 units in hidden layer are optimal number of hidden layers that is verified by the DTREG commercial software [68]. So the total weights become 40. Total 69 correctly segmented masses are used for classification where 25% images are used for training and 75% are used for testing purpose and the overall classification for benign is 83.87% and for malignant 90.91% that is shown in table 6.3. Three examples are shown as an output of the proposed ANN-based classifier in figure 6.18, 6.19 and 6.20. In figure 6.18 and 6.19, masses are correctly classified as benign and malignant respectively whereas in figure 6.20 original mass is malignant but it is classified as benign using the proposed method.

The classification performance can be assessed in terms of the sensitivity and specificity of the system. Sensitivity (SN) is the proportion of actual positives which are correctly identified and it is mathematically defined in equation 6.29 and specificity (SP) is the proportion of negatives which are

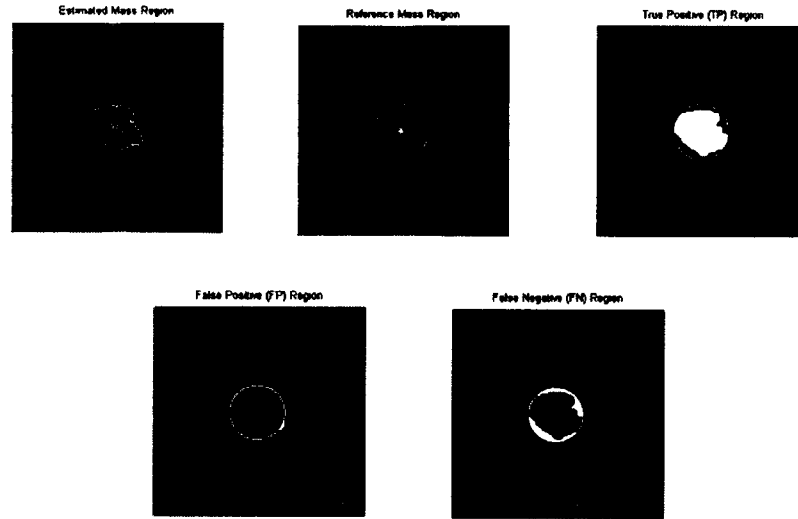


Figure 6.16: Segmentation Performance Evaluation- Example 1

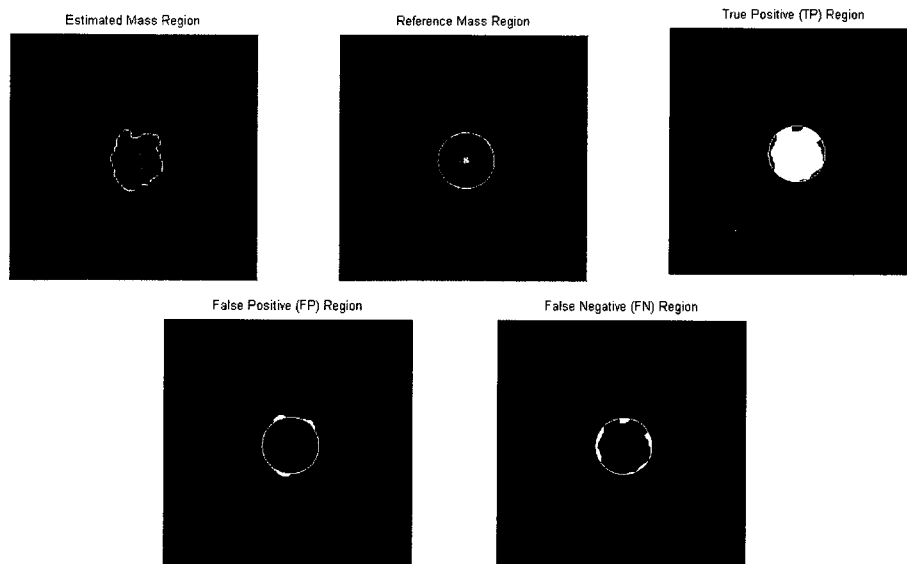


Figure 6.17: Segmentation Performance Evaluation- Example 2

Table 6.2: Proposed ANN Structure for Mass Classification

Input Units	n	7
Hidden Units	$(n + 1) * \frac{2}{3}$	5 (verified by DTREG [68])
Output Unit	1	1
Weights	$n * ((n + 1) * \frac{2}{3}) + (n + 1) * \frac{2}{3}$	40

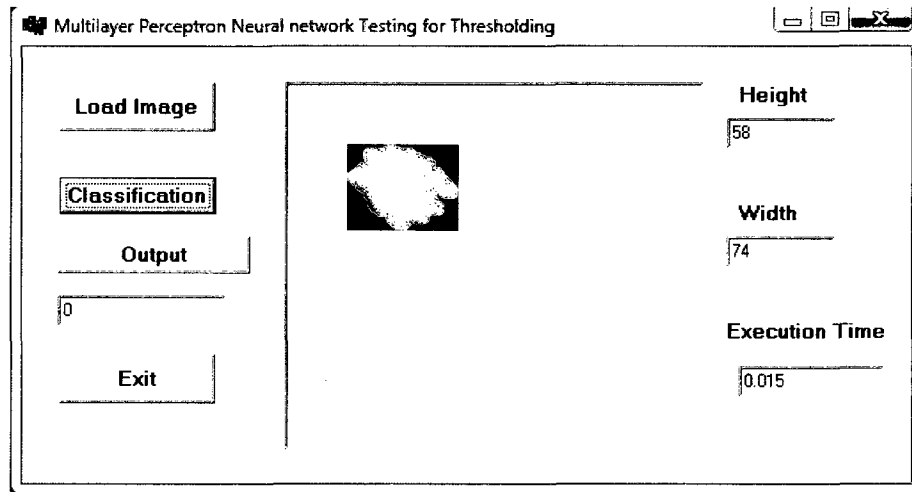


Figure 6.18: Mass Classification- Example 1, Correct Classification (Benign mass)

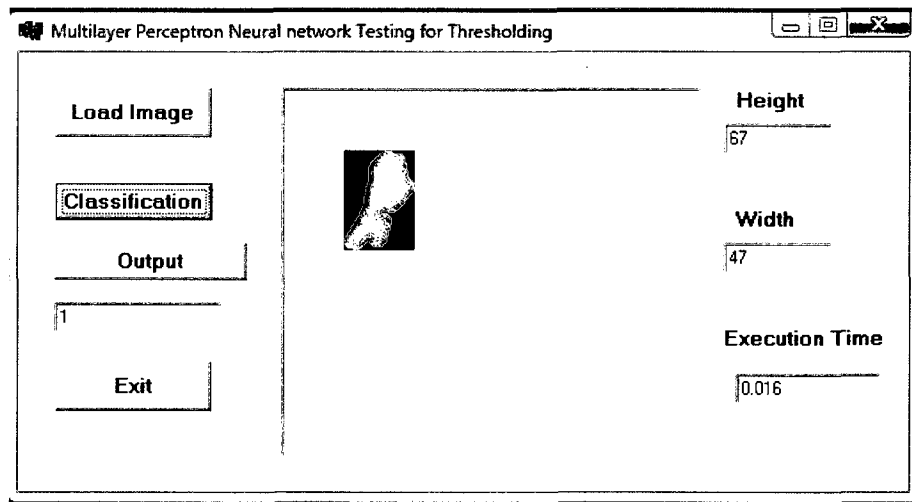


Figure 6.19: Mass Classification- Example 2, Correct Classification (Malignant mass)

Table 6.3: Comparative Statement of Classification

Correct Classification (%)		Misclassification (%)		Radiologist Misclassification (%)	
Benign	Malignant	Benign	Malignant	Benign	Malignant
83.87	90.91	16.63	9.09	65-90	Not Available

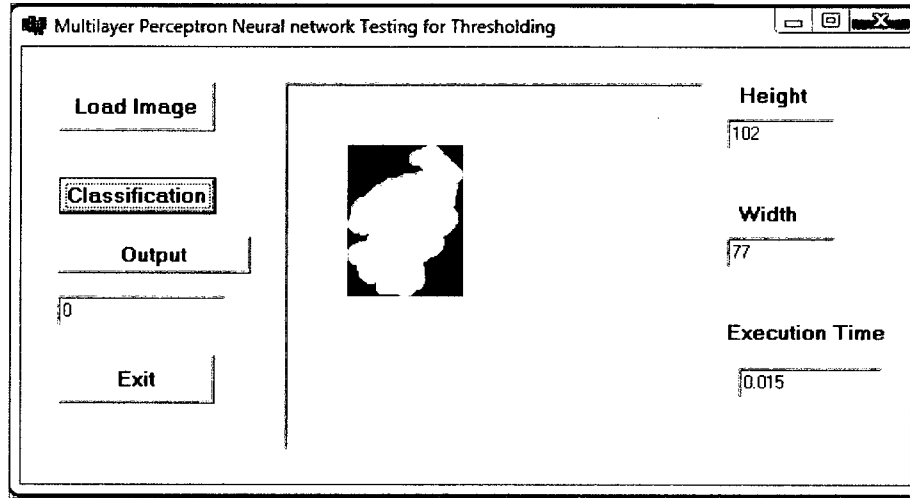


Figure 6.20: Mass Classification- Example 3, Misclassification (Classified as Benign, actually Malignant)

correctly identified and is mathematically defined in equation 6.30. The sensitivity and specificity of the proposed system is shown in table 6.4 and it is compared to the radiologist's sensitivity.

$$SN = \frac{TP}{TP + FN} \quad (6.29)$$

$$SP = \frac{TN}{TN + FP} \quad (6.30)$$

Table 6.4: Comparative Statement of Sensitivity of the Proposed System to the Radiologist

TP	TN	FP	FN	Sensitivity (%)	Specificity (%)	Radiologist's Sensitivity (%)
20	26	5	2	90.91	83.87	75

6.6 Validation

To validate the proposed segmentation algorithm an artificially blurred maple leaf image is used. Maple leaf image is combined into the artificially generated background. Then the proposed image enhancement (equation 6.28) and automatic seeded region growing method is applied to the enhanced image and finally the contour of the maple leaf image and the maple leaf itself is extracted from the original image. In example 1, figure 6.21, 6.22 and 6.23 shows the complete output using the proposed method from object, background to enhanced image, automated seed selection, segmented image, extracted contour of the maple leaf and extracted maple leaf.

In example 2, we combined two maple leaves in one image with very small gray level difference. In that case the proposed system can detect and extract only one leaf. This is the limitation of the proposed method so far. However, the proposed method was developed assuming that only one mass belongs to one ROI. But if there are two masses in one ROI very close to each other, in that case the proposed method can detect only one. Figure 6.24, 6.25 and 6.26 shows the output of leaf extraction using proposed method for the example 2. It is observed that in the figure 6.25, the seed points suppose to be in the block 18 and 45. However there is a slight gray level difference between them and because of this the Sum Average for block 18 is 1 whereas for block 45 it is 0.9994 and finally the block 18 was selected as the seed block.

In example 3, we choose the background the breast tissue where the tissue is inhomogeneous. Then the scaled maple leaf image was combined with the inhomogeneous background. The combined image is then passed through the proposed image enhancement method followed by block division, automated seed selection, segmentation and finally the maple leaf extraction. The proposed method works fine in inhomogeneous background and the results are shown in figure 6.27, 6.28 and 6.29.

6.7 Conclusion

In this chapter a computer-aided system for detection, segmentation and classification of masses is developed. Initially the x-ray label is removed using global Otsu thresholding technique followed by connected component labeling. Pectoral muscle is removed using automatic region growing method. ROI is extracted using peak analysis from the histogram of the breast tissue and contrast enhancement of the ROI is done using nonlinear operator. Seed selection is the major challenge in any region growing method. Automated seeded region growing is proposed for image segmentation. Automated seed selection is done using Haralick texture features. Sum Average is found the most discriminative features among 13 features and finally segmented image is being smoothed using mathematical mor-

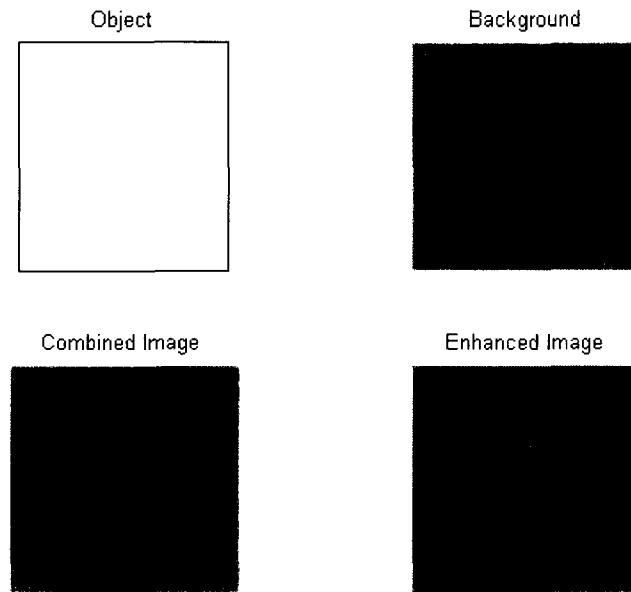


Figure 6.21: Validation- Example 1, Background, Object, Combined Image and Enhanced Image

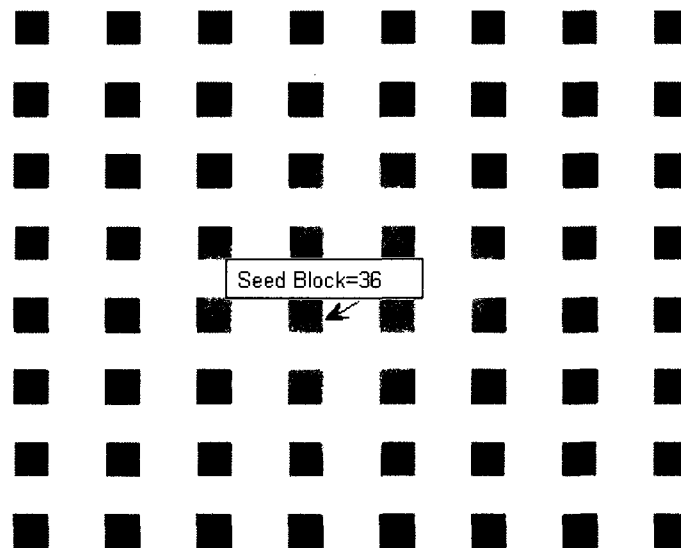


Figure 6.22: Validation- Example 1, Automated Seed Selection

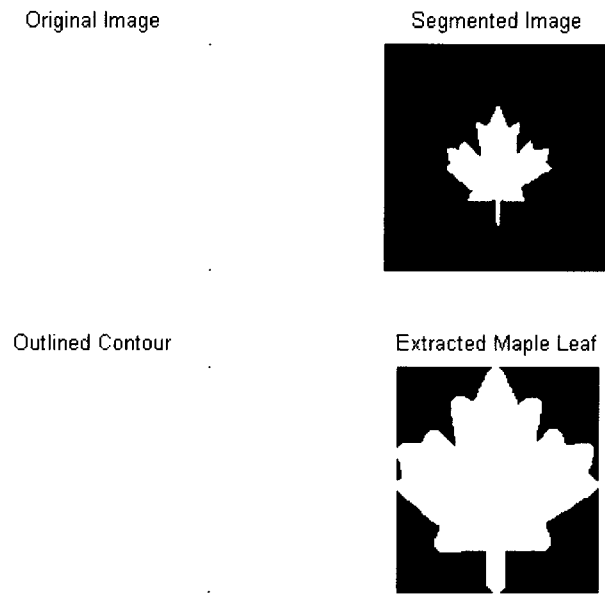


Figure 6.23: Validation- Example 1, Segmented Leaf, Extracted Contour and Extracted Leaf

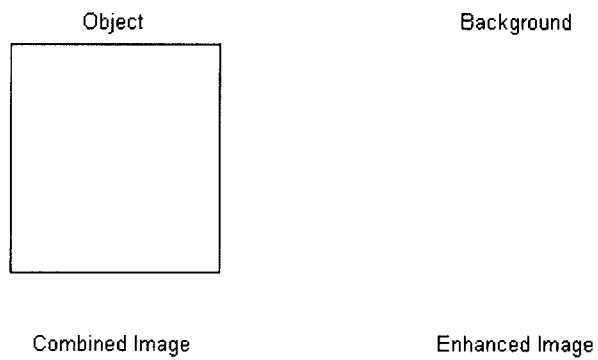


Figure 6.24: Validation- Example 2, Background, Object, Combined Image and Enhanced Image

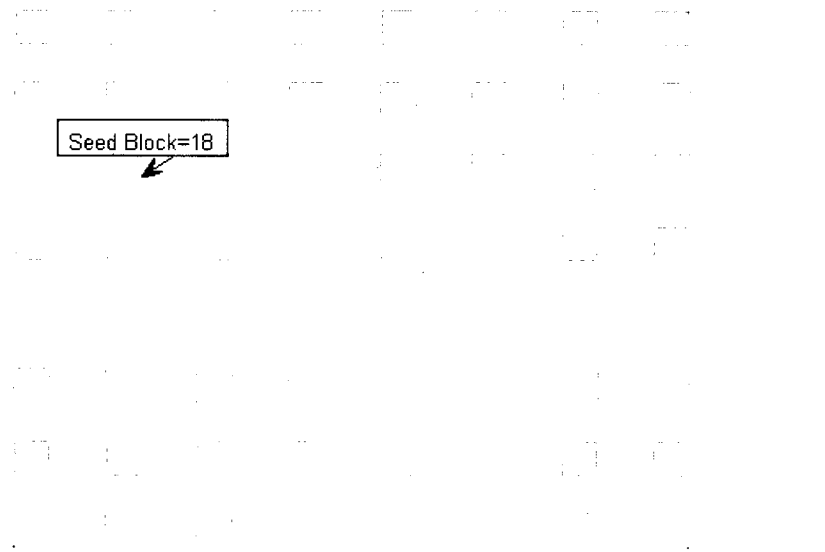


Figure 6.25: Validation- Example 2, Automated Seed Selection

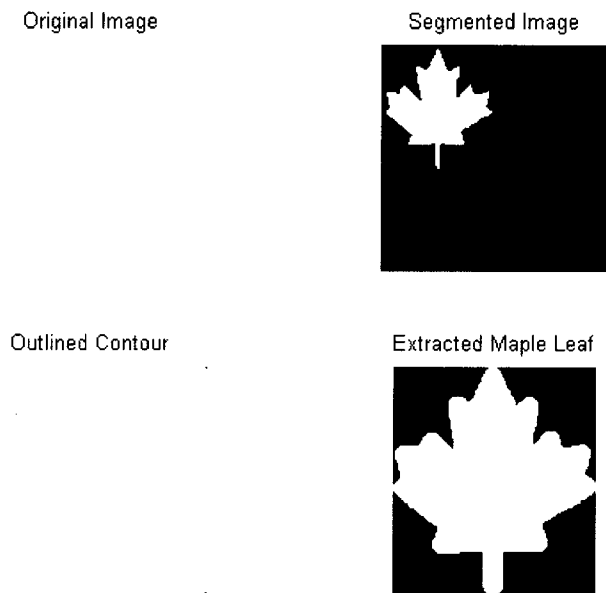


Figure 6.26: Validation- Example 2, Segmented Leaf, Extracted Contour and Extracted Leaf

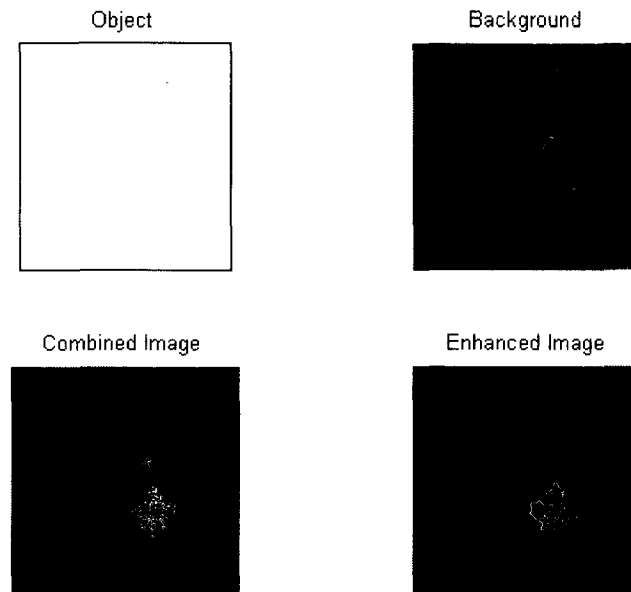


Figure 6.27: Validation- Example 3, Background, Object, Combined Image and Enhanced Image

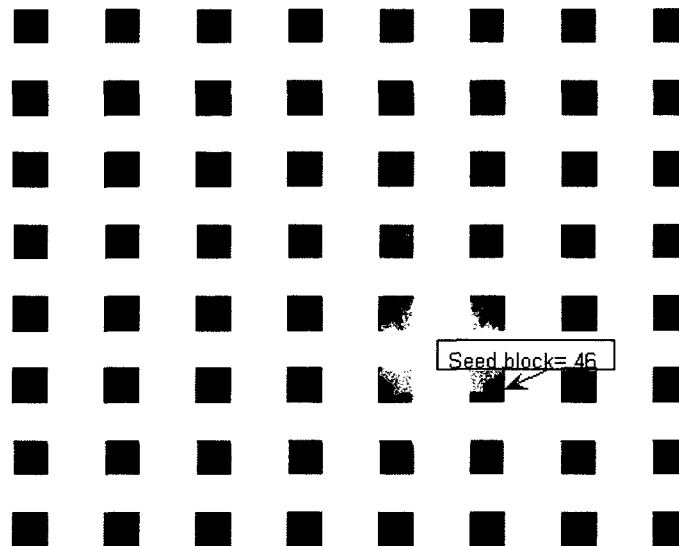


Figure 6.28: Validation- Example 3, Automated Seed Selection

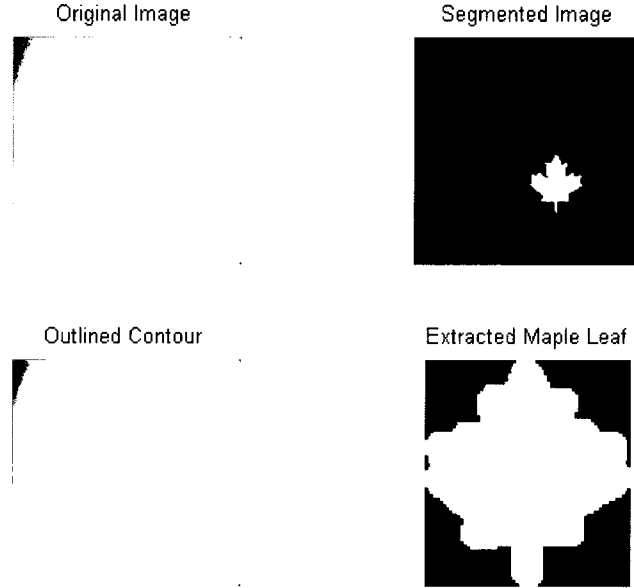


Figure 6.29: Validation- Example 3, Segmented Leaf, Extracted Contour and Extracted Leaf

phology operators. Performance of the proposed method is evaluated using two quantitative features Completeness (CM) and Correctness (CR). Correct segmentation is achieved 84.15% that is very much promising. To validate the proposed segmentation algorithm an artificially blurred maple leaf image was used and it was successfully recovered from the blurred image. Then two maple leaves were added to the background and the proposed system is failed to extract the second maple leaf. Finally, original inhomogeneous breast tissue was used as a background of the maple leaf and the proposed system successfully recovered the maple leaf. An artificial neural network of 3 layers is proposed for mass classification. Using the proposed ANN classifier, 90.91% sensitivity and 83.87% specificity is achieved which is very much encouraging compare to the radiologist's sensitivity 75%.

Chapter 7

Conclusion and Future Research

7.1 Conclusions

In this thesis an efficient ANN based thresholding technique is developed based on the work reported in references [7, 63, 67] to binarize the document image with non-uniform and complex background. To achieve that an exhaustive search was conducted to find the optimal parameters like optimal features and window size using grid-based method. This is a local pixel's characteristics based binarization method. Eight statistical and textural features reported in [67] were used for that purpose and these features are obtained from the neighborhoods around every pixel. It was calculated that total of 255 subsets can be generated without repetition using these eight features. These 255 subsets were placed in the left column of the grid and in the right columns, different window size like 3x3, 5x5, 7x7 and 9x9 were placed. For each subset four binary images is obtained corresponding to each window size and passed them through commercial OCR ABBYY. The obtained character recognition rate for each binary image is tabulated in the grid. Finally the subset and window size that intersects at highest recognition rate is termed as optimal features and optimal window size. Performance of the optimal features were evaluated in terms of PSNR and computational time and it was found that **pixel value**, **mean** and **entropy** and window size 3x3 are the most significant features subset in document recognition application. The optimal features subset was validated using known base documents and watermark images. The proposed method successfully binarized the watermarked images in almost all cases except one where the watermark is very much

complex. Finally, using the proposed efficient ANN-based thresholding method, 99.25% correct recognition rate is achieved for a database of about 15000 characters. The result demonstrates better performance compared to the methods described in the literature. The proposed method requires fewer features resulting in less computational cost and makes it suitable for document recognition application.

ANN is also used in computer-aided mass detection and classification from digitized mammograms in biomedical imaging applications. In this thesis, a method was presented for preprocessing mammograms for x-ray label removal, pectoral muscle removal, peak analysis of the histogram which results in obtaining the ROI. The ROI is then enhanced using the proposed nonlinear operator. The enhanced ROI is used as an input to the proposed automated seed region growing method to extract the mass from the ROI. In any region growing method the main challenge is the seed selection. This is done here using Haralick texture features [98]. Haralick texture features were used for the first time in this application to mass extraction. The ROI is divided into 32x32 non-overlapped regions and then 13 Haralick features were calculated for each block. Sum Average is found to be the most significant feature that easily discriminates mass and non-mass region. The image block with maximum sum average represents the seed block and maximum gray level of that block is the seed point. Region growing starts from that point and grows iteratively resulting in extracted mass. The extracted mass is then smoothed using mathematical morphological operators dilation and erosion. Performance of the proposed segmentation is evaluated using two quantitative measures CM and CR. For 10 ROIs, using the proposed method the 78% CM and 94% CR achieved. MiniMIAS database was used in this application where a mass is marked as a circle in the above database which is not the case in reality. That's why completeness becomes less in this simulation. 84.15% correct segmentation is achieved in this application using the proposed method. To validate the proposed method, artificially blurred maple leaf image is used. The maple leaf is successfully recovered from the low contrast image using the proposed method. The proposed method however, failed in one situation in recovering the maple leaf. In the next experiment, two maple leaves were embedded in the background and finally one maple leaf was recovered successfully. Therefore, in case if there are more than one mass present in the ROI, the proposed system can detect only one of them. In another example, the maple leaf was superimposed on a background of real breast with inhomogeneous tissue. The proposed method successfully recovered the maple leaf. The statistical and textural features are then utilized to develop the ANN-based classifier to classify the mass into benign or malignant. Seven features were proposed in this application and they are used as input to the ANN-based classifier. The proposed ANN consists of three layers. The first layer (input) has

seven neurons, the hidden layer has five neurons while the output layer has only one neuron. Output value of 0 means benign while 1 represents malignant. The proposed system produces 90.91% sensitivity and 83.87% specificity whereas the radiologist's sensitivity is around 75%. The computational time to classify a mass using the proposed system requires around 50 ms. The proposed system shows promising results both in terms of recognition rate and computational time.

The main disadvantage of the ANN-based local thresholding method is its computational cost, which hampers its implementation in real world applications. ANN-based local thresholding method was investigated in quality inspection of two parts gelatin capsule in pharmaceutical manufactured industry applications. The main challenge in this application was to extract the clear, transparent capsule from the background. An edge-based binarization method was proposed for quality inspection using Sobel edge detector. The developed image processing system meets the industry requirements of inspecting 1000 capsules per minutes with an acceptable accuracy of detecting defects over 95% of size 0.2 mm and larger. In addition to the edge-based method, border tracing and least square circle fitting was utilized to approximate both body and cap dome to a circle. The radius of these circles initially identifies dents on the capsule. The proposed system is very flexible and it can be adjusted based on the industry requirements to accept the different size bubbles, dirt and other foreign body inside the capsule. The proposed system provided 97.45% sensitivity and 95.67% specificity which is very much promising.

7.2 Contributions and Future work

An efficient ANN-based local thresholding method was developed based on MLP NN to extract the characters from the non-uniform and complex background using local pixels characteristics like statistical and textural features. A method is devised to find the optimal parameters like optimal features and optimal window size. An exhaustive search was conducted for that purpose. The proposed method using the optimal parameters outperformed other local or global thresholding method discussed in the literature.

Automated seeded region growing method is developed to extract mass from the enhanced ROI using Haralick texture features. Nonlinear operator is proposed to enhance the contrast of the ROI. Haralick texture features were used first time to extract mass from the ROI so far and it provides very promising output in this application. ANN classifier based on the statistical and textural features were proposed for mass classification and 90.91% sensitivity was achieved from the proposed system.

Edge-based binarization along with border tracing and circle fitting is proposed to develop a

computer vision-based quality controller in pharmaceutical industry application. It is rarely found any system can work on a capsule where both body and cap are transparent which is the case in this application. The proposed system is flexible and meets the industry requirements of inspecting 1000 capsules per minute.

In future works, it is important to determine the potential of new ultrasound computer-aided diagnostic techniques to detect and diagnose breast cancer. The x-ray mammogram has several disadvantages. Mostly it is not very much effective for women with dense breasts. Moreover, it involves radiation, which makes it undesirable. Also, some patients feel discomfort and pain during the procedure. An alternative diagnostic tool to X-ray mammography would be of value and ultrasound imaging can be one solution. It can represent muscle and soft tissue very well as well. It delivers live image, no side effects and widely available. The expected system should achieve high sensitivity and specificity for small breast cancers and microcalcifications particularly in cases not suitable for mammography, as well as reduce the number of breast biopsies across Canada and the world.

References

- [1] R.E. Bird, "Professional quality assurance for mammography screening programs," Intl. J. Radiol., Vol. 175, pp. 587-605, 1990.
- [2] Neural Network- <http://www.doc.ic.ac.uk/~nd/surprise.96/journal/vol4/cs11/report.html>.
- [3] J.A. Anderson, "An introduction to neural networks," MIT Press, Cambridge, MA, 1995.
- [4] T. Kohonen, "Self-organization and associative memory," Springer-Verlag, Berlin, 1984.
- [5] J.J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," Proc. of the National Academy of Sciences of the USA, Vol. 79 no. 8 pp. 2554-2558, April 1982.
- [6] M.A. Sid-Ahmed, "Image processing theory, algorithms and architectures," 1st ed., McGraw-Hill, New York, 1995.
- [7] Y. Alginahi, "Computer analysis of composite documents with non-uniform background," PhD Thesis, Electrical and Computer Engineering, University of Windsor, Windsor, ON, 2004.
- [8] M. Sezgin and B. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation," Journal of Electronic Imaging, Vol. 13, No. 1, 2004.
- [9] P.K. Sahoo, S. Soltani, A.K.C. Wong, "A Survey of thresholding techniques," Computer Vision, Graphics and Image Processing, Vol. 41, pp. 233-260, 1988.
- [10] S.U. Le, S.Y. Chung, and R.H. Park, "A comparative performance study of several global thresholding techniques for segmentation," Graph. Models Image Process. Vol. 52, pp. 171-190, 1990.
- [11] L.G. Shapiro and G.C. Stockman, "Computer Vision," pp. 279-325, New Jersey, Prentice-Hall, 2001.
- [12] R. Ohlander, K. Price, and D.R. Reddy, "Picture segmentation using a recursive region splitting method, Computer Graphics and Image Processing, Vol. 8, pp. 313-333, 1978.
- [13] M.I. Sezan, "A peak detection algorithm and its application to histogram-based image data reduction," CVGIP. Vol. 29, pp. 47-59, 1985.
- [14] S.C. Sahasrabudhe and K.S.D. Gupta, "A valley-seeking threshold selection technique," Computer Vision and Image Processing, (A. Rosenfeld, L.Shapiro eds.), pp. 55-65, Academic Press, 1992.

-
- [15] J.C. Ohio, "Automatic threshold selection using the wavelet transform," *CVGIP*, Vol. 56, pp. 205-218, 1994.
 - [16] A. Rosenfeld, P. De laTorre, "Histogram concavity analysis as an aid in threshold selection," *IEEE SMC*, Vol. 13, pp. 231-235, 1983.
 - [17] J. Weszka, A. Rosenfeld, "Histogram modification for threshold selection," *IEEE SMC*, Vol. 9, pp. 38-52, 1979.
 - [18] P. Rosin, "Unimodal thresholding," *Pattern Recognition*, Vol. 34, no. 11, pp. 2083-2096, 2001.
 - [19] W. Tsai, "Moment-preserving thresholding," *Comp. Graphics and Image Processing*, Vol. 29, pp. 377-393, 1985.
 - [20] Y. Solihan and C.G. Leedham, "Integral ratio: A new class of global thresholding techniques for handwriting images," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 21, no. 8, pp. 761-768, 1999.
 - [21] G. Leedham, S. Varma, A. Patankar and V. Govindaraju, "Separating text and background in degraded document images- A comparison of global thresholding techniques for multi-stage thresholding," *Proc. 8th Int. Workshop in Handwriting Recognition*, pp. 244-249, 2002.
 - [22] M.J. Carlotto, "Histogram analysis using a scale-space approach," *IEEE PAMI*, Vol. 9, pp. 121-129, 1987.
 - [23] G. Sapiro, V. Casselles, "Histogram modification via partial differential equations," *Proc. of ICIP 95*, pp. 632-635, 1995.
 - [24] A Fletcher, K. Kasturi, "A robust algorithm for text string separation from mixed text /graphics images," *IEEE Trans. Pattern Anal, Machine Intell, PAMI-10*, pp. 910-918, 1988.
 - [25] H. J. Trussel, "Comments on picture thresholding using iterative selection method," *IEEE Trans. Syst. Man Cybern. SMC-9*, pp. 311-319, 1979.
 - [26] M. K. Yanni and E. Horne, "A new approach to dynamic thresholding," *EUSIPCO94: 9th European Conf. Sig. Process.* Vol. 1, pp. 34-44, 1994.
 - [27] Y. Yang and H. Yan, "An adaptive logical method for binarization of degraded document images," *Pattern Recogn.* Vol. 33, pp. 787-807, 2000.
 - [28] B.R. Lee, K. Park, H. Kang, H. Kim, and C. Kim, "Adaptive local binarization method for recognition of vehicle license plates," *LNCS-Combinatorial Image Analysis*, Springer Berlin, Vol. 3322, pp. 646-655, 2005.
 - [29] T. W. Ridler and S. Calvard, "Picture thresholding using an iterative selection method," *IEEE Trans. Syst. Man Cybern. SMC*, Vol. 8, pp. 630-632, 1978.
 - [30] C. K. Leung and F. K. Lam, "Performance analysis of a class of iterative image thresholding algorithms," *Pattern Recogn.* Vol. 29, No. 9, pp. 1523-1530, 1996.
 - [31] R.C. Gonzalez and R.E. Woods, "Digital image processing," 2nd ed., Pearson Education Asia, 2004.
 - [32] S. Huang, M. Ahmadi, M.A. Sid-Ahmed, "A hidden Markov model-based character extraction method," *Pattern Recognition*, Vol. 41, pp. 2890-2900, 2008.
-

-
- [33] J.M. Prager, "Extracting and labeling boundary segments in natural scenes," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 2, no. 1, pp. 16-27, 1980.
 - [34] W.A. Perkins, "Area segmentation of images using edge points," *IEEE Trans. on Pattern Recognition and Machine Intelligence*, Vol. 2, no. 1, pp. 8-15, 1980.
 - [35] H. Yalcyn, I. Bozma and M. Tezol, "BUVIS-A novel approach to real-time automated visual inspection," *Proc. of Texas Instruments DSP Challenge- Top 20 European Papers*, pp. 9-15, 1998.
 - [36] J.N. Kapur, P.K. Sahoo, and A.K.C. Wong, "A new method for gray-level picture thresholding using the entropy of the histogram," *Graphical Models and Image Processing*, Vol. 29, pp. 273-285, 1985.
 - [37] N. Otsu, "A threshold selection method from gray level histograms," *IEEE Trans. On Systems, Man and Cybernetics*, Vol. 9, pp. 62-66, 1979.
 - [38] J. Kittler, J. Illingworth, "Minimum error thresholding," *Pattern Recognition*, Vol. 19, pp. 41-47, 1986.
 - [39] J. Kittler. And J. Illingworth, "On Threshold Selection Using clustering Criteria," *IEEE SMC*, Vol.15, pp. 652-655, 1985.
 - [40] C.K. Chow and T. Kaneko, "Automatic detection of the left ventricle from cineangiograms," *Comp. and Biomed. Research*, Vol. 5, pp. 388-410, 1972.
 - [41] Y. Nakagawa and A. Rosenfeld, "Some experiments on variable thresholding," *Pattern Recognition*, Vol. 11, no. 3, pp. 191-204, 1979.
 - [42] D.E. Lloyd, "Automatic target classification using moment invariant of image shapes," *Technical Report, RAE IDN AW126*, Farnborough-UK, December 1985.
 - [43] S. Cho, R. Haralick, and S. Yi, "Improvement of Kittler and Illingworthss minimum error thresholding," *Pattern Recognition*, Vol. 22, pp. 609-617, 1989.
 - [44] T.W. Riddler, S. Calvard, "Picture thresholding using an iterative thresholding selection method," *IEEE Trans. SMC*, Vol. 8, pp. 630-632, 1978.
 - [45] M.A. Sid-Ahmed, "A hardware structure for the automatic selection of multi-level thresholds in digital images," *Pattern Recognition*, Vol. 25, no. 12, pp. 1517-1528, 1992.
 - [46] C.V. Jawahar, P.K. Biswas, and A.K. Ray, "Investigations on fuzzy thresholding based on fuzzy clustering," *Pattern Recogn.* Vol. 30, No. 10, pp. 1605-1613, 1997.
 - [47] A.S. Abutaleb, "Automatic thresholding of gray-level pictures using two-dimensional entropy," *Computer Vision, Graphics and Image Processing*, Vol. 47, pp. 22-32, 1989.
 - [48] T. Pun, "A new method for grey-level picture threshold using the entropy of the histogram," *EUFL4SIP:Signal Processing*, Vol. 2, No. 3, pp. 223-237, July 1980.
 - [49] T. Pun, "Entropic thresholding, A new approach," *Computer Graphics and Image Processing*, Vol. 16, pp. 210-239, 1981.
 - [50] A. D. Brink, and N.E. Pendock, "Minimum cross entropy threshold selection," *Pattern Recognition*, Vol. 29, pp. 179-188, 1996.
-

-
- [51] C.H. Li, C.K. Lee, "Minimum cross-entropy thresholding," *Pattern Recognition*, Vol. 26, pp. 617-625, 1993.
- [52] A.G. Shanbag, 'Utilization of Information Measure as a Means of Image Thresholding,' *CVGIP*, Vol. 56, pp. 414-419, 1994.
- [53] P. Sahoo, C. Wilkins, and J. Yeager, "Threshold selection using Renyis entropy," *Pattern Recognition*, Vol. 30, pp. 71-84, 1997.
- [54] W. Niblack, "An introduction to image processing," pp. 115-116, Prentice-Hall, Englewood Cliffs, NJ 1986.
- [55] J. Sauvola and M. Pietaksinen, "Adaptive document image binarization," *Pattern Recogn.* Vol. 33, pp. 225-236, 2000.
- [56] S. D. Yanowitz and A. M. Bruckstein, "A new method for image segmentation," *Comput. Graph. Image Process.* Vol. 46, pp. 82-95, 1989.
- [57] D. Shen and H. H. S. Ip, "A Hopfield neural network for adaptive image segmentation: An active surface paradigm," *Pattern Recogn. Lett.* Vol. 18, pp. 37-48, 1997.
- [58] J.M. White, G.D. Rohrer, "Image segmentation for optical character recognition and other applications requiring character image extraction," *IBM J. Res. Dev.* Vol. 27, No. 4, pp. 400-411, 1983.
- [59] H. Yan and J. Wu, "Character and line extraction from color map images using a multi-layer neural network," *Pattern Recognition Letters*, 15 pp. 97-103, 1994.
- [60] R. Koker and Y. Sari, "Neural Network based automatic threshold selection for an industrial vision system," *Proc. Int. Conf. on Signal Processing*, pp. 523-525, 2003.
- [61] A. Khashman, B. Sekeroglu, "Document image binarization using a supervised neural network," *International Journal of Neural Systems*, Vol. 18, No. 5, pp. 405-418, 2008.
- [62] N. Papamarkos, "A technique for fuzzy document binarization," *Proc. of the ACM Symposium on Document Engineering*, pp. 152-156, 2001.
- [63] Y. Alginahi, "Thresholding and character recognition in security documents with watermarked background," *Proc. Int. Conf. on Digital Image Computing: Techniques and Applications (DICTA'08)*, pp. 220-225, 2008 .
- [64] P.L. Rosin, and E. Ioannidis, "Evaluation of global image thresholding for change detection," *Pattern Recognition Letters*, Vol. 24, pp. 2345-2356, 2003.
- [65] S. Fisher, "Digital image processing: Skewing and thresholding," MSc. Thesis, University of New South Wales, Sydney, Australia, 2000.
- [66] O.D Trier and A.K. Jain, "Goal-directed evaluation of binarization methods," *IEEE Trans. on Pattern Recognition and Machine Intelligence*, Vol. 17, No. 12, pp. 1191-1201, 1995.
- [67] Y. Alginahi, M.A. Sid-Ahmed and M. Ahmadi, "Local thresholding of composite documents using Multi-layer Perceptron Neural Network," *The 47-th IEEE International Midwest Symposium on Circuits and Systems*, pp. 209-212, 2004.
- [68] <http://www.dtrek.com/>
-

-
- [69] Y.J. Zhang, "A survey on evaluation methods for image segmentation," *Pattern Recognition* Vol. 29, pp. 1335-1346, 1996.
- [70] J. He, Q.D.M. Do, A.C. Downton and J.H. Kim, "A comparison of binarization methods for historical archive documents," *Proc. of the 2005 Eight Intl. Conf. on Document Analysis and Recognition*, Vol. 1, pp. 538-542, 2005.
- [71] A.C. Karloff, N.E. Scott, and R. Muscedere, "A flexible design for a cost effective, high throughput inspection system for pharmaceutical capsules," in *IEEE International Conference Industrial Technology (ICIT'2008)*, Chengdu, China, April 21-24, 2008.
- [72] Daiichi Jitsugyo Viswill Co. Ltd., "Capsule Visual Inspection System - CVIS-SXX-E," http://www.viswill.jp/English/CVIS_E/cvis_index_e.html, 2005.
- [73] M.J. Islam, M. Ahmadi and M.A. Sid-Ahmed, "Image processing techniques for quality inspection of gelatin capsules in pharmaceutical applications," in *10th International Conference on Control, Automation, Robotics and Vision (ICARCV 2008)*, Hanoi, Vietnam, December 17-22, 2008.
- [74] M.J. Islam, Q.M.J. Wu, M. Ahmadi and M.A. Sid-Ahmed, "Grey scale image segmentation using Minimum Error Thresholding techniques," in *6th International Conference for Upcoming Engineers (ICUE' 2007)*, Ryerson University, Toronto, ON, Canada, pp. 1-5 (CD Version), May 28-29, 2007.
- [75] A. Chakraborty, L.H. Staib and J.S. Duncan, "Deformable boundary finding influenced by region homogeneity," *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 624-627, Seattle, WA, June 1994.
- [76] M. Sharafi, M. Fathy and M.T. Mahmoudi, "A classified and comparative study of edge detection algorithms," *International Conference on Information Technology: Coding and Computing*, pp. 117-120, April 8-10, 2002.
- [77] D.H. Ballard and C. Brown, "Computer vision," Prentice Hall, 1982.
- [78] A. Chottera and M. Shridhar, "Feature extraction of manufactured parts in the presence of spurious surface reflections," *Canadian Journal of Electrical and Computer Engineering*, Vol. 7, no. 4, pp. 29-33, 1982.
- [79] Border tracing, digital image processing lectures- The University of Iowa, <http://www.icaen.uiowa.edu/dip/LECTURE/Segmentation2.html#tracing>.
- [80] C.J. Baines, D.V. McFarlane and A.B. Miller, "The role of the reference radiologist: estimates of inter-observer agreement and potential delay in cancer detection in the national screening study," *Investiga Radiology*, Vol. 25, pp. 971-976, 1990.
- [81] J.A. Harvey, L.L. Fajardo, C.A. Inis, "Previous mammograms in patients with impalpable breast carcinomas: retrospective vs. blinded interpretation," *American J. Radiol.* Vol. 161, No. 6, pp. 1167-1172, 1993.
- [82] S.C. Yang, C.M. Wany et.al., "A Computer-aided system for mass detection and classification in digitized mammograms," *J. Bio. Med. Engg.- Appl., Basis and Comm.*, Vol. 17, pp. 215-228, 2005.
-

-
- [83] V. Gimenez, D. Manrique, J. Rios, A. Vilarrasa, "Iterative method for automatic detection of masses in digital mammograms for computer-aided diagnosis," In Proceedings of SPIE The International Society for Optical Engineering 3661, Vol. 2, pp. 1086-1093, 1999.
- [84] T. Matsubara, H. Fujita et.al., "Development of new schemes for detection and analysis of mammographic masses," *Intell. Inform. Sys.*, pp. 63-66, 1997.
- [85] M. Abdel-Mottaleb, C.S. Carman, C.R. Hill, S. Vafai, "Locating the boundary between breast skin edge and the background in digitized mammograms," *Digital Mammography*, Elsevier, Amsterdam, pp. 467-470, 1996.
- [86] Y.D. Li, M. Kallergi, L.P. Clarke, V.K. Jain, R.A. Clark, "Markov random field for tumor detection in digital mammography," *IEEE Trans. Med. Imaging*, Vol. 14, No. 3, pp. 565-576, 1995.
- [87] M. Kallergi, K. Woods, L.P. Clarke, W. Qian, R.A. Clark, "Image segmentation in digital mammography: comparison of local thresholding and region growing algorithms," *Comput. Med. Imaging Graph.* Vol. 16, No. 5, pp. 231-323, 1992.
- [88] S.M Lai, X. Li, W.F. Biscof, "On techniques for detecting circumscribed masses in mammo-grams," *IEEE Trans. Med. Imaging*, Vol. 18, No.4, pp. 377-386, 1989.
- [89] S. Morrison, L.M. Linnett, "A model based approach to object detection in digital mammogra-phy," in Proceedings of the IEEE International Conference on Image Processing, Vol. 2, Kobe, Japan, pp. 182-186, 1999.
- [90] G.D. Tourassi, R. Vargas-Voracek, "Computer-assisted detection of mammographic masses: a template matching scheme based on mutual information," *Med. Phys.* Vol. 30, pp. 2123-2130, 2003.
- [91] A. Mencattini, G. Rabottino, M. Salmeri, R. Lojacono, E. Colini, "Breast mass segmenta-tion in mammographic images by an effective region growing algorithm," in LNCS, Springer, Heidelberg, Vol. 5259, pp. 948-957, 2008.
- [92] S.A. Hojjatoleslami, J. Kittler, "Region growing: a new approach," *IEEE Trans. Image Process.* Vol. 7, no. 7, pp. 1079-1084, 1998.
- [93] A.J. Mendez, P.G. Tahoces, M.J. Lado, M. Souto, J.J. Vidal, "Computer-aided diagnosis: automatic detection of malignant masses in digitized mammograms," *Med. Phys.* Vol. 25, No. 6, pp. 957-964, 1998.
- [94] M. Sameti, R.K. Ward, "A fuzzy segmentation algorithm for mammogram partitioning," *Digital Mammography*, Elsevier, Amsterdam, pp. 471-474, 1996.
- [95] K. Wongsritong, K. Kittayaruasiriwat, F. Cheevasuvit, K. Dejhan, A. Somboonkaew, "Contrast enhancement using multi-peak histogram equalization with brightness preserving," in IEEE Asia-Pacific Conference on Circuits and Systems Proceedings, pp. 455-458, 1998.
- [96] P. Undrill, R. Gupta, S. Henry, M. Dowing, "Texture analysis and boundary refinement to outline mammography masses," in Proceedings of the 1996 IEE colloquium on Digital mam-mography, pp. 51-56, 1996.
- [97] N. Petrick, H.P. Chan, B. Sahiner, B. Helvie, "Combined adaptive enhancement and region growing segmentation of breast masses on digitized mammograms," *Med. Phys.* Vol. 26, pp. 1642-1654, 1999.
-

-
- [98] R.M. Haralick, K. Shanmugam, I.K. Denstein, "Textural features for image classification," IEEE Transactions on Systems, Man and Cybernetics, Vol. 3, No. 6, pp. 610-621, 1973.
 - [99] R. O. Duda, P. E. Hart, D. G. Stork, "Pattern classification," John Wiley and Sons, second edition, 2001.
 - [100] H.P. Chan, D. Wei, M.A. Helvie, B. Sahiner et.al., "Computer-aided classification of mammographic masses and normal tissue: linear discriminant analysis in texture feature space," J. Phy. Med. Biol., Vol. 40, pp. 857-876, 1995.
 - [101] B. Sahiner, H.P. Chan, M.A. Helvie, M.M. Goodsitt, "Computerized classification of benign and malignant masses on digitized mammograms: a study of robustness," Acad. Radiology. Vol. 7, pp. 1077-1084, 2000.
 - [102] G.M. Brake, N. Karssemeijer, J.H.C.L. Hendriks, "An automatic method to discriminate malignant masses from normal tissue in digital mammograms," Phys. Med. Biol. Vol. 45, no. 10, pp. 2843-2857, 2000.
 - [103] P.J.G. Lisboa, "A review of evidence of health benefits from artificial neural networks in medical intervention," Neural Networks, Vol. 15, pp. 11-39, 2000.
 - [104] H.D. Cheng, X.J. Shi, L.M. Hu, X.P. Cai, H.N. Du, "Approaches for automated detection and classification of masses in mammograms," Pattern Recognition, Vol. 39, pp. 646-668, 2006.
 - [105] R. M. Haralick, S. R. Sternberg, X. Zhuang, "Image analysis using mathematical morphology," IEEE Trans. Pattern Anal. Machine Intell. PAMI-9, pp. 532-550, 1987.
 - [106] D. Raba, A. Oliver, J. Marti, M. Peracaula, J. Espunya, "Breast mass segmentation with pectoral muscle suppression on digital mammograms," LNCS, Springer-Verlag Berlin, Vol. 3523, pp. 471-478, 2005
 - [107] S.M. Kwok, R. Chandrashekhar, Y. Attikiouzel, "Automatic pectoral muscle segmentation on mediolateral oblique view mammograms," IEEE Trans. on Med. Imag., Vol. 23, pp. 1129-1140, 2004.
 - [108] J. Suckling et al., "The Mammographic Image Analysis Society Digital Mammogram Database Excerpta Medica," International Congress Series, Vol. 1069, pp. 375-378, 1994.
-

Appendix A

Recognition Rate- Grid-based Method

This table represents the recognition rate of a sample document image (figure 4.2) using a commercial OCR ABBYY 7.0. For 8 features 255 different subsets are available. Each subset is used to obtain the binary image at window size 3x3, 5x5, 7x7 and 9x9. The binary image is then passed through the commercial OCR ABBYY 7.0 and the obtained recognition rate is tabulated in the following tables A.1, A.2 and A.3. Therefore, total $255 \times 4 = 1020$ binary images as well as 1020 recognition rates are available for one document image.

Features					Features				
Recognition Rate					Recognition Rate				
Subset	3x3	5x5	7x7	9x9	Subset	3x3	5x5	7x7	9x9
{1}	0	0	0	0	{1,4,8}	74.47	74.01	73.24	72.48
{2}	0	0	0	0	{1,5,6}	99.24	99.54	99.08	98.93
{3}	0	0	0	0	{1,5,7}	98.32	98.47	98.01	96.79
{4}	0	0	0	0	{1,5,8}	94.95	95.41	94.65	94.8
{5}	0	0	0	0	{1,6,7}	82.42	84.71	83.94	81.65
{6}	0	0	0	0	{1,6,8}	83.18	81.65	79.36	78.59
{7}	0	0	0	0	{1,7,8}	84.71	83.94	82.42	81.65
{8}	0	0	0	0	{2,3,4}	93.88	89.3	87.77	87
{1,2}	98.93	99.24	99.39	98.47	{2,3,5}	75.54	77.06	76.3	75.54
{1,3}	95.72	95.11	95.87	95.26	{2,3,6}	89.3	90.83	92.35	91.59
{1,4}	95.26	94.65	94.95	94.5	{2,3,7}	83.18	84.71	82.42	81.65
{1,5}	98.78	99.08	99.24	98.93	{2,3,8}	87.77	89.3	90.06	86.24
{1,6}	98.93	99.08	98.62	98.47	{2,4,5}	85.47	87.77	83.94	81.65
{1,7}	94.8	94.04	95.11	95.26	{2,4,6}	86.24	84.74	87.47	85.62
{1,8}	92.35	93.88	94.5	94.04	{2,4,7}	83.18	81.65	80.12	82.42
{2,3}	95.11	95.87	94.8	94.65	{2,4,8}	77.83	77.06	75.54	74.01
{2,4}	94.5	95.41	95.26	95.11	{2,5,6}	0	0	0	0
{2,5}	0	0	0	0	{2,5,7}	91.59	92.35	93.88	90.06
{2,6}	0	0	0	0	{2,5,8}	90.83	89.3	90.06	88.53
{2,7}	91.59	92.35	94.04	93.88	{2,6,7}	94.04	93.88	95.11	91.59
{2,8}	80.12	81.65	83.94	82.42	{2,6,8}	91.59	90.83	91.13	92.05
{3,4}	90.06	89.3	84.71	88.53	{2,7,8}	77.83	77.06	76.30	77.22
{3,5}	85.47	84.71	86.24	83.18	{3,4,5}	75.54	74.01	72.48	73.24
{3,6}	84.71	86.24	83.94	85.47	{3,4,6}	77.06	75.54	76.30	76.76
{3,7}	86.24	87	87.77	85.47	{3,4,7}	89.60	89.3	89.14	88.99
{3,8}	82.42	87.77	83.94	85.47	{3,4,8}	83.18	81.65	83.94	83.18
{4,5}	61.77	69.42	64.07	66.36	{3,5,6}	64.83	66.36	65.60	63.30
{4,6}	62.54	69.36	69.42	67.13	{3,5,7}	69.42	69.89	70.18	70.03
{4,7}	65.6	67.89	68.65	67.13	{3,5,8}	77.22	77.06	76.30	75.23
{4,8}	70.18	68.65	66.36	69.42	{3,6,7}	74.77	74.01	74.31	75.23
{5,6}	0	0	0	0	{3,6,8}	71.71	72.48	74.01	73.70
{5,7}	83.18	84.71	88.53	85.47	{3,7,8}	72.48	70.95	72.17	71.25
{5,8}	83.94	81.65	82.42	83.18	{4,5,6}	70.95	69.42	68.81	69.11
{6,7}	88.53	86.24	85.47	84.71	{4,5,7}	85.02	84.71	83.94	85.47
{6,8}	71.71	77.06	75.54	74.77	{4,5,8}	78.59	77.06	77.83	77.06
{7,8}	77.83	75.54	77.06	76.3	{4,6,7}	72.48	70.95	71.71	72.17
{1,2,3}	96.79	95.41	96.64	95.87	{4,6,8}	74.16	72.48	72.17	72.78
{1,2,4}	96.18	94.95	95.87	95.41	{4,7,8}	69.11	69.42	68.65	70.18
{1,2,5}	100	99.85	99.85	99.69	{5,6,7}	85.32	84.71	85.47	85.02
{1,2,6}	99.24	99.54	99.08	98.01	{5,6,8}	76.30	77.06	78.29	75.54
{1,2,7}	96.18	95.11	95.72	93.88	{5,7,8}	72.17	72.48	73.24	70.95
{1,2,8}	95.41	95.26	96.18	95.72	{6,7,8}	69.72	69.42	71.71	68.65
{1,3,4}	96.02	94.95	96.18	95.87	{1,2,3,4}	94.50	93.88	94.19	93.12
{1,3,5}	95.57	96.18	96.02	96.18	{1,2,3,5}	93.12	92.35	92.66	91.59
{1,3,6}	94.34	93.88	92.35	90.83	{1,2,3,6}	93.58	94.04	93.12	93.58
{1,3,7}	94.19	92.35	91.59	90.83	{1,2,3,7}	95.41	94.5	93.88	93.58
{1,3,8}	80.12	84.71	83.94	82.42	{1,2,3,8}	95.11	94.65	94.19	95.41
{1,4,5}	98.32	99.24	98.93	98.32	{1,2,4,5}	95.11	95.72	95.41	94.95
{1,4,6}	74.01	77.06	78.59	77.83	{1,2,4,6}	91.59	92.35	92.05	91.13
{1,4,7}	76.03	75.54	74.01	74.47	{1,2,4,7}	92.35	91.59	92.51	91.13

Figure A.1: Recognition rate in various window size- Table A.1

Features	Recognition Rate				Features	Recognition Rate			
Subset	3x3	5x5	7x7	9x9	Subset	3x3	5x5	7x7	9x9
{1,2,4,8}	98.01	98.32	98.47	97.71	{3,4,5,8}	70.95	69.42	69.11	67.89
{1,2,5,6}	99.24	99.39	99.54	98.78	{3,4,6,7}	64.83	66.36	67.89	68.04
{1,2,5,7}	98.17	99.08	98.47	97.71	{3,4,6,8}	69.42	67.13	64.83	63.30
{1,2,5,8}	97.71	97.55	97.25	98.01	{3,4,7,8}	70.18	70.95	69.42	67.89
{1,2,6,7}	97.71	98.47	98.01	98.32	{3,5,6,7}	67.89	68.65	66.36	67.13
{1,2,6,8}	98.78	98.17	97.71	97.40	{3,5,6,8}	74.01	72.48	69.42	73.24
{1,2,7,8}	99.24	98.78	98.47	97.55	{3,5,7,8}	69.42	74.01	66.36	64.83
{1,3,4,5}	96.94	96.18	95.41	94.65	{3,6,7,8}	72.48	69.42	70.95	67.89
{1,3,4,6}	96.18	95.72	96.64	95.87	{4,5,6,7}	68.20	66.36	67.89	65.60
{1,3,4,7}	96.18	95.41	95.72	94.65	{4,5,6,8}	70.18	67.89	69.42	66.36
{1,3,4,8}	96.64	96.02	95.41	94.19	{4,5,7,8}	68.65	69.11	67.89	68.20
{1,3,5,6}	92.35	93.88	93.12	90.83	{4,6,7,8}	66.36	64.83	61.77	63.30
{1,3,5,7}	94.19	94.65	95.11	94.50	{5,6,7,8}	70.95	69.42	67.89	67.13
{1,3,5,8}	92.35	92.66	93.12	92.51	{1,2,3,4,5}	91.59	90.83	90.06	88.99
{1,3,6,7}	93.88	93.12	92.66	92.35	{1,2,3,4,6}	90.06	89.30	91.59	89.76
{1,3,6,8}	91.59	92.35	90.06	92.05	{1,2,3,4,7}	89.45	90.06	89.30	90.67
{1,3,7,8}	92.35	91.59	91.13	90.83	{1,2,3,4,8}	93.43	91.59	92.05	90.98
{1,4,5,6}	95.11	94.65	95.41	94.19	{1,2,3,5,6}	90.83	87.77	89.30	88.53
{1,4,5,7}	91.59	92.35	92.05	93.12	{1,2,3,5,7}	89.30	87.00	88.53	89.45
{1,4,5,8}	93.58	93.12	92.66	93.43	{1,2,3,5,8}	90.83	91.59	93.12	92.51
{1,4,6,7}	91.59	90.83	91.13	90.06	{1,2,3,6,7}	88.53	87.92	89.30	89.76
{1,4,6,8}	91.90	92.66	90.83	91.28	{1,2,3,6,8}	87.77	86.24	87.00	87.31
{1,4,7,8}	92.05	90.83	90.06	89.76	{1,2,3,7,8}	89.76	88.84	90.98	90.06
{1,5,6,7}	93.12	92.35	92.05	91.28	{1,2,4,5,6}	94.19	93.12	92.51	92.05
{1,5,6,8}	90.83	91.28	91.59	90.67	{1,2,4,5,7}	92.05	91.59	90.98	92.51
{1,5,7,8}	88.53	89.30	88.99	88.07	{1,2,4,5,8}	88.07	87.77	88.99	88.53
{1,6,7,8}	86.54	86.24	85.93	85.47	{1,2,4,6,7}	87.77	86.24	85.78	86.70
{2,3,4,5}	68.65	69.42	67.89	66.36	{1,2,4,6,8}	88.07	87.16	87.77	86.39
{2,3,4,6}	74.01	72.48	70.95	69.42	{1,2,4,7,8}	87.00	86.39	88.38	87.92
{2,3,4,7}	77.83	77.06	76.30	74.77	{1,2,5,6,7}	91.74	91.59	90.83	91.13
{2,3,4,8}	74.77	74.01	72.48	73.39	{1,2,5,6,8}	94.19	93.88	93.12	93.58
{2,3,5,6}	76.76	75.54	74.31	73.24	{1,2,5,7,8}	90.83	90.06	88.38	91.28
{2,3,5,7}	64.83	68.65	63.30	61.77	{1,2,6,7,8}	86.24	84.71	86.70	85.63
{2,3,5,8}	63.30	68.20	68.81	64.83	{1,3,4,5,6}	90.52	90.06	88.53	88.99
{2,3,6,7}	69.42	72.94	72.48	68.81	{1,3,4,5,7}	92.97	91.44	89.76	90.52
{2,3,6,8}	74.01	72.17	71.10	70.64	{1,3,4,5,8}	88.53	86.70	87.77	87.31
{2,3,7,8}	72.48	71.71	70.95	69.42	{1,3,4,6,7}	90.83	89.76	89.30	88.53
{2,4,5,6}	69.42	71.10	72.48	71.71	{1,3,4,6,8}	87.77	86.24	87.00	87.16
{2,4,5,7}	70.95	69.27	71.71	73.24	{1,3,4,7,8}	87.00	85.93	85.47	85.17
{2,4,5,8}	67.89	66.06	68.65	69.42	{1,3,5,6,7}	90.06	89.60	88.99	88.53
{2,4,6,7}	70.95	68.65	67.89	67.13	{1,3,5,6,8}	88.53	87.46	88.38	87.46
{2,4,6,8}	71.71	67.89	69.42	70.95	{1,3,5,7,8}	90.06	88.84	87.92	87.46
{2,4,7,8}	61.77	65.60	63.30	61.01	{1,3,6,7,8}	92.51	91.13	90.06	90.52
{2,5,6,7}	64.83	67.43	69.42	68.65	{1,4,5,6,7}	86.39	84.71	85.78	85.47
{2,5,6,8}	66.36	67.13	67.89	70.18	{1,4,5,6,8}	89.76	88.23	88.53	88.99
{2,5,7,8}	67.13	64.83	68.65	63.30	{1,4,5,7,8}	87.00	86.24	84.71	85.47
{2,6,7,8}	66.36	65.29	64.83	64.07	{1,4,6,7,8}	90.98	89.60	90.52	90.06
{3,4,5,6}	69.42	66.36	67.89	64.83	{1,5,6,7,8}	85.78	85.47	86.24	84.86
{3,4,5,7}	73.24	72.48	70.95	70.18	{2,3,4,5,6}	74.01	77.06	72.48	74.01
{1,4,7}	76.03	75.54	74.01	74.47	{1,2,4,7}	92.35	91.59	92.51	91.13

Figure A.2: Recognition rate in various window size- Table A.2

Features	Recognition Rate				Features	Recognition Rate			
Subset	3x3	5x5	7x7	9x9	Subset	3x3	5x5	7x7	9x9
{2,3,4,5,7}	75.23	74.01	73.24	72.48	{1,2,3,6,7,8}	81.65	78.29	80.12	79.20
{2,3,4,5,8}	77.83	77.06	75.23	73.70	{1,2,4,5,6,7}	72.78	71.41	70.18	73.09
{2,3,4,6,7}	74.01	75.54	72.48	76.15	{1,2,4,5,6,8}	68.65	70.95	70.03	71.41
{2,3,4,6,8}	72.48	70.95	71.41	70.18	{1,2,4,5,7,8}	74.77	73.55	74.16	75.69
{2,3,4,7,8}	67.89	69.42	70.18	70.80	{1,2,4,6,7,8}	76.15	74.92	74.01	73.85
{2,3,5,6,7}	74.01	71.71	69.42	68.81	{1,2,5,6,7,8}	80.28	78.75	78.29	79.36
{2,3,5,6,8}	71.41	72.17	72.48	69.72	{1,3,4,5,6,7}	77.83	77.06	76.30	76.61
{2,3,5,7,8}	75.54	73.24	71.56	72.94	{1,3,4,5,6,8}	69.42	70.95	68.65	67.89
{2,3,6,7,8}	66.36	69.27	68.04	70.34	{1,3,4,5,7,8}	68.35	69.42	68.04	67.13
{2,4,5,6,7}	68.65	66.36	70.95	70.64	{1,3,4,6,7,8}	66.36	68.96	67.89	68.65
{2,4,5,6,8}	64.83	67.89	64.07	67.13	{1,3,5,6,7,8}	67.89	70.03	70.34	71.41
{2,4,5,7,8}	72.48	70.18	71.25	69.57	{1,4,5,6,7,8}	72.78	72.02	74.46	70.64
{2,4,6,7,8}	70.64	69.42	67.89	66.06	{2,3,4,5,6,7}	64.83	66.36	67.89	68.65
{2,5,6,7,8}	74.77	71.71	71.41	70.18	{2,3,4,5,6,8}	70.03	67.89	68.65	67.13
{3,4,5,6,7}	70.95	69.42	71.41	70.18	{2,3,4,5,7,8}	77.83	76.15	74.77	77.22
{3,4,5,6,8}	64.83	67.89	66.36	68.65	{2,3,4,6,7,8}	73.55	74.46	73.09	72.17
{3,4,5,7,8}	77.06	74.16	75.69	72.63	{2,3,5,6,7,8}	72.78	72.02	71.10	70.64
{3,4,6,7,8}	74.77	72.94	74.01	72.78	{2,4,5,6,7,8}	67.89	69.42	66.06	67.58
{3,5,6,7,8}	71.41	69.72	67.89	66.36	{3,4,5,6,7,8}	71.71	74.16	73.39	71.41
{4,5,6,7,8}	64.83	67.89	66.36	68.35	{1,2,3,4,5,6,7}	90.83	90.06	90.52	88.84
{1,2,3,4,5,6}	80.12	77.06	75.54	74.77	{1,2,3,4,5,6,8}	87.46	86.24	86.70	85.93
{1,2,3,4,5,7}	74.77	74.01	72.48	70.95	{1,2,3,4,5,7,8}	93.12	92.35	91.44	90.98
{1,2,3,4,5,8}	76.30	74.77	73.09	72.48	{1,2,3,4,6,7,8}	92.66	91.59	90.67	90.52
{1,2,3,4,6,7}	74.01	70.64	72.48	69.88	{1,2,3,5,6,7,8}	87.46	86.70	85.78	85.47
{1,2,3,4,6,8}	73.24	71.41	72.94	72.17	{1,2,4,5,6,7,8}	88.99	88.38	87.92	87.00
{1,2,3,4,7,8}	76.15	75.54	74.46	73.24	{1,3,4,5,6,7,8}	90.21	88.99	87.46	86.09
{1,2,3,5,6,7}	80.12	77.06	75.54	73.70	{2,3,4,5,6,7,8}	69.42	70.95	66.36	69.88
{1,2,3,5,6,8}	74.77	74.01	73.39	72.32	{1,2,3,4,5,6,7,8}	90.06	88.53	86.70	89.45
{1,2,3,5,7,8}	75.54	76.15	74.01	74.77					

Figure A.3: Recognition rate in various window size- Table A.3

VITA AUCTORIS

Mohammed Jahirul Islam was born in Comilla, Bangladesh, in June, 1975. He received his BSc. (Hon's) and MSc. degree in electronics and computer science in 1996 and 1997 respectively from Shahjalal University of Science and Technology, Sylhet, Bangladesh and his Master of Applied Science (MAsc.) degree in electrical engineering in 2003 from the Ryerson University, Toronto, ON, Canada. He is currently a candidate in the electrical and computer engineering Ph.D. program at the University of Windsor. His research interests include machine learning and computer vision, image processing, image segmentation, document image analysis, mammogram analysis and breast cancer detection.